

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 15:34:28 Search time 457.59 Seconds  
(without alignments)  
2352.554 Million cell updates/sec

Title: US-09-824-647-16\_COPY\_1\_627

Perfect score: 627

Sequence: 1 cgcaggcagaccattggac.....cagtgccctgtccagctcg 627

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

N\_Geneseq\_032802:\*  
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	627	100.0	2095	20 AAV82825	Human GP88 autocri
2	627	100.0	2124	14 AAQ56794	Granulin coding se
3	627	100.0	2198	14 AAQ49052	Granulin coding se
4	627	100.0	2432	21 AAC78180	Human cancer assoc
5	615	98.1	1779	12 AAQ14339	Human epithelin pr
6	406.6	64.8	1767	12 AAQ14338	Rat epithelin prec
7	400.8	63.9	2137	20 AAV82824	Mouse GP88 autocri
8	390.4	62.3	1767	12 AAQ14340	Mouse epithelin pr
9	381.2	60.8	561	21 AAA69780	Human ovarian carc

10	297.4	47.4	757	22 AAH07141	Human cDNA clone (
11	297.4	47.4	1630	22 AAH16370	Human cDNA sequenc
12	59	9.4	341	12 AAQ14953	Chicken epithelin
13	50.2	8.0	51	21 AAC15693	Human secreted pro
14	48.6	7.8	539	12 AAQ14952	Bovine epithelin p
15	41.6	6.6	3729	23 AAS94337	DNA encoding novel
16	40.2	6.4	10732	21 AAQ10594	Gene encoding a su
17	40	6.4	1824	23 AAS81488	DNA encoding novel
18	40	6.4	2243	22 ABA08657	Human 'extensin hom
19	40	6.4	2772	23 AAS72787	DNA encoding novel
20	40	6.4	2772	23 AAS90525	DNA encoding novel
21	40	6.4	2850	23 AAS79895	DNA encoding novel
22	39.2	6.3	799	19 AAV55831	Nucleotide sequenc
23	39.2	6.3	1925	20 AAX50924	Epstein Barr Virus
24	39.2	6.3	1926	21 AAS50254	Epstein Barr virus
25	39.2	6.3	1926	22 AAF82902	EBV tethering prot
26	39.2	6.3	2580	21 AAX75454	Nucleotide sequenc
27	39.2	6.3	5452	20 AAX90923	Anti-sense strand
28	39.2	6.3	8705	20 AAZ23778	Vector pShuttle DN
29	39.2	6.3	9600	19 AAV21683	Nucleotide sequenc
30	39.2	6.3	10380	20 AAZ22248	Plasmid pCisEBON f
31	39.2	6.3	10596	14 AAQ51731	Plasmid pCisEBON f
32	39.2	6.3	10596	17 AAT40348	Nucleotide sequenc
33	39.2	6.3	10596	20 AAX15850	DNA clone pCEK Cl.
34	39.2	6.3	16080	21 AAS59553	Micromonospora DNA
35	38.4	6.1	109519	22 AAS08693	DNA encoding novel
36	38	6.1	1267	23 AAS67932	DNA encoding novel
37	38	6.1	1833	23 AAS67931	Murine transport a
38	37.8	6.0	565	22 AAK53492	Lysine decarboxyla
39	37	5.9	3269	17 AAT34583	Human brain expres
40	36.8	5.9	233	22 AAK18681	Aspergillus niger
41	36.8	5.9	2028	15 AAQ73737	Aspergillus niger
42	36.8	5.9	2476	15 AAQ73736	Myobacterium tube
43	36.6	5.8	4403765	22 AAI99683	P. putida KT2440-a
44	36.6	5.8	4411529	22 AAI99682	
45	36.2	5.8	1680	22 AAF61077	

#### ALIGNMENTS

#### RESULT 1

AAV82825

ID AAV82825 standard; cDNA; 2095 BP.

XX

XX

AC AAV82825;

XX 15-MAR-1999 (first entry)

DT

XX Human GP88 autocrine growth factor cDNA.

DE

XX GP88; granulin; epithelin; human; growth factor; autocrine; tumour;

KW cancer; viral infection; antagonist; therapy; diagnosis; ss.

XX Homo sapiens.

OS

PH Key Location/Qualifiers

FT CDS 13..1794

FT /\*tag= a

XX

PN WO9852607-A1.

XX

PD 26-NOV-1998.

XX

PF 22-MAY-1998; 98WO-US10555.

XX

PR 16-DEC-1997; 97US-0991862.

PR 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

PA

XX Mouse GP88 autocri

XX Mouse epithelin pr

XX Serrero G;

DR WPI: 1999-045276/04.  
DR P-PSDB; AAW85475.  
XX  
PT Composition containing antagonist of growth factor GP88 - useful for  
PT treating cancer and viral diseases and also for diagnosing disease  
PT from altered GP88 expression  
XX  
PS Disclosure; Fig 9A; 86pp; English.  
XX  
CC This human cDNA sequence includes a coding region for GP88 (see  
CC AAW85474), an 88 kDa glycoprotein autocrine growth factor and  
CC epithelin/granulin precursor that is expressed in a tightly  
CC regulated manner in normal cells, is overexpressed and unregulated  
CC in highly tumorigenic cells derived from normal cells, and which  
CC acts as a stringently required growth stimulator for the  
CC tumorigenic cells. Inhibition of GP88 expression or action in the  
CC tumorigenic cells results in an inhibition of the tumorigenic  
CC properties of the overproducing cells. Antagonists to GP88 are  
CC used to treat diseases associated with increased expression of  
CC GP88, particularly cancer but also viral infections. Fragments of  
CC GP88 are used to raise specific antibodies (used as antagonists,  
CC as diagnostic reagents and for delivering toxins or other  
CC compounds to GP88-expressing cells) and to screen for antibodies.  
CC Antisense oligonucleotides can also be used as antagonists.  
CC Methods are provided for diagnosing disease, or determining  
CC susceptibility to disease, resulting from altered GP88 activity.  
XX  
SQ Sequence 2095 BP; 364 A; 681 C; 625 G; 425 T; 0 other;  
Query Match 100.0%; Score 627; DB 20; Length 2095;  
Best Local Similarity 100.0%; Pred. No. 1.6e-154;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 cgcaggcagaccattggaacctggtgagctgggtggccttaacagcagggtggtgct 60  
DB 1 cgcaggcagaccattggaacctggtgagctgggtggccttaacagcagggtggtgct 60  
QY 61 ggaacgcggtgccagatggtcagttctgacctgctgacctgctgacctgacctgacct 120  
DB 61 ggaacgcggtgccagatggtcagttctgacctgctgacctgctgacctgacctgacct 120  
QY 121 gccagctacagctgctgcctgctgctgctgctgctgctgctgctgctgctgctgct 180  
DB 121 gccagctacagctgctgcctgctgctgctgctgctgctgctgctgctgctgctgct 180  
QY 181 ctgggtggccctgccagtgatgccactgctgctgctgctgctgctgctgctgctgct 240  
DB 181 ctgggtggccctgccagtgatgccactgctgctgctgctgctgctgctgctgctgct 240  
QY 241 gtcacaggacttcagttgctgctgctgctgctgctgctgctgctgctgctgctgct 300  
DB 241 gtcacaggacttcagttgctgctgctgctgctgctgctgctgctgctgctgctgct 300  
QY 301 cactgctgccacgggcttcactgctgctgctgctgctgctgctgctgctgctgctgct 360  
DB 301 cactgctgccacgggcttcactgctgctgctgctgctgctgctgctgctgctgctgct 360  
QY 361 ggtacaaactcgtgggtgctgctgctgctgctgctgctgctgctgctgctgctgct 420  
DB 361 ggtacaaactcgtgggtgctgctgctgctgctgctgctgctgctgctgctgctgct 420  
QY 421 tcacagtgctggttatggtgctgctgctgctgctgctgctgctgctgctgctgctgct 480  
DB 421 tcacagtgctggttatggtgctgctgctgctgctgctgctgctgctgctgctgctgct 480  
QY 481 tctgtgaagacaggtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 540  
DB 481 tctgtgaagacaggtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 540  
QY 541 cgtctgcatcacaccacaggggacccacccctgccaagaagctccctgccaagaggact 600  
DB 541 cgtctgcatcacaccacaggggacccacccctgccaagaagctccctgccaagaggact 600

QY 601 aacagggcagtgccctgtccagctcg 627  
DB 601 aacagggcagtgccctgtccagctcg 627  
RESULT 2  
AAQ56794  
ID AAQ56794 standard; DNA; 2124 BP.  
XX  
AC AAQ56794;  
XX  
DT 22-APR-1994 (first entry)  
XX  
DE Granulin coding sequence.  
XX  
KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
KW granulocytes; leucocytes; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
XX misc\_difference 1379..1381  
XX /\*tag- a  
XX /\*transl\_except- ATG encodes Valine.  
XX misc\_difference 1640..1642  
XX /\*tag- b  
XX /\*transl\_except- CAG encodes Glycine.  
XX  
PN W09315195-A.  
XX  
PD 05-AUG-1993.  
XX  
PF 28-FEB-1992; 92WO-CA00089.  
XX  
PR 03-FEB-1992; 92US-0829233.  
XX  
PA (SOLO/) SOLOMON S.  
XX  
PI Solomon S;  
XX  
DR WPI: 1993-320328/40.  
DR P-PSDB; AAR48673.  
XX  
PT New cystine rich granulin peptide(s) from leucocyte(s) - are  
PT keratinocyte inhibitors useful topically for wound healing  
XX  
PS Disclosure; Figure 4c; 53pp; English.  
XX  
CC The granulin inhibits keratinocytes and is useful in formulations  
CC for promoting the healing of wounds.  
XX  
SQ Sequence 2124 BP; 383 A; 685 C; 630 G; 426 T; 0 other;  
Query Match 100.0%; Score 627; DB 14; Length 2124;  
Best Local Similarity 100.0%; Pred. No. 1.6e-154;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 cgcaggcagaccattggaacctggtgagctgggtggccttaacagcagggtggtgct 60  
DB 14 cgcaggcagaccattggaacctggtgagctgggtggccttaacagcagggtggtgct 73  
QY 61 ggaacgcggtgccagatggtcagttctgacctgctgacctgctgacctgacctgacct 120  
DB 74 ggaacgcggtgccagatggtcagttctgacctgctgacctgctgacctgacctgacct 133  
QY 121 gccagctacagctgctgcctgctgctgctgctgctgctgctgctgctgctgctgct 180  
DB 134 gccagctacagctgctgcctgctgctgctgctgctgctgctgctgctgctgctgct 193  
QY 181 ctgggtggccctgccaggttgatgccactgctgctgctgctgctgctgctgctgctgct 240

Db	194	ctgggtggccctgccaggttgatgccactgctgcggcgccactctctgcacatctttacc	253
QY	241	gtctcagggaactccagttgctgcccttcccagagcgctgcatcggggatggccat	300
Db	254	gtctcagggaactccagttgctgcccttcccagagcgctgcatcggggatggccat	313
QY	301	caactgtgtccacggggcttcactgcagtcagacggcgcatctctctccaaagatca	360
Db	314	caactgtgtccacggggcttcactgcagtcagacggcgcatctctctccaaagatca	373
QY	361	ggtaacaactccgtgggtgccatccagtcacctgatgtcagttcgaatgccggacttc	420
Db	374	ggtaacaactccgtgggtgccatccagtcacctgatgtcagttcgaatgccggacttc	433
QY	421	tcacagctgtgttatggctcgtatggctctgggggtgctgcccatgcccaggcttc	480
Db	434	tcacagctgtgttatggctcgtatggctctgggggtgctgcccatgcccaggcttc	493
QY	481	tgtgtgaagacaggggtgcactgctgtccgcacggtgcctctgcgacctgttcacacc	540
Db	494	tgtgtgaagacaggggtgcactgctgtccgcacggtgcctctgcgacctgttcacacc	553
QY	541	cgctgcatcacaccaccaggggacccaccctctggcaagaagctccctgcccagaggact	600
Db	554	cgctgcatcacaccaccaggggacccaccctctggcaagaagctccctgcccagaggact	613
QY	601	aacaggcgagtgccctgttccagctcg	627
Db	614	aacaggcgagtgccctgttccagctcg	640

RESULT 3  
AAQ49052  
ID AAQ49052 standard; DNA: 2198 BP.

AA  
AC AAO49052:

22-APR-1994 (first entry)

XX DE Granulin coding sequence.

**XX** Granulin: keratinocytes: wound healing; inhibition; peptide;  
**KW** Granulin: keratinocytes: wound healing; inhibition; peptide;

KW granulocytes: leucocytes: ss::  
KW granulocytes: leucocytes: ss::

XX Homo sapiens.

XX	Key	Location/Qualifiers
FH		

FH	key	Location/Qualifiers
FT	misc difference	1453..1455

```
FT misc_difference 1433..1433
FT /*taq= a
```

```
FT /-tag= a
FT /transl_except= ATG encodes Valine.
```

```
FI      /crust_except= 110
FT      misc_difference 1714..1716
```

```
FT difference 1/14..1/10
FT /*tag= a
```

```
FT. /transl_except= CAG encodes Glycine.
```

XX  
/XXXXXXXXXXXX/

PN W09315195-A..

XX

PD 05-AUG-1993.

[illegible]

PF 28-FEB-1992; 92WO-CA00089.

XX  
00-0000-1000  
0000-00000000

PR 03-FEB-1992; 92US-0829233.

XX  
FM  
'COTCA' COTOMON S

PA (SOLO/) SOLOMON S.

XX pr Solomon S.

PI Solomon S;  
YY

XX  
DR WPT: 1993-320328/40.

DR WPI; 1993-320328/40.  
DR P-PSDB: AAR48673.

DR P-FSDB; AAK48073.  
XX

AA  
PT  
New cystine rich granulin peptide(s) from leucocyte(s)

PT keratinocyte inhibitors useful topically for wound healing

XX  
XX  
FI keratinocyte tumor cells used for copycatting for wound

PS Disclosure; Figure 4c; 53pp; English.

**THE UNIVERSITY OF CHICAGO**

[illegible]

## 4 RESULT

RESULT  
AAC78180

AAC/8180  
ID AAC78180 standard; cDNA; 2432 BP.

XX  
ID AAC/0100

AAC78180;

AC  
XX  
FAC/0100,

DT 08-FEB-2001 (first entry)

XX

DE Human cancer associated gene sequence SEQ ID NO:574.

Human; cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytostatic; proliferative; vulnery; immunomodulator; antidiabetic; antialstemic; antirheumatic; antithyritic; antiviral; antinflammatory; antithyroid; antiallergic; antibacterial; cardiac; dermatologic; neuroprotective; thrombolytic; coagulant; nootropic; vasotropic; antipsoriasis; antiangiogenic; gene therapy; inflammation; immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW	allergic reaction; graft versus host disease; organ rejection;
KW	haemostatic; thrombolytic; cardiovascular disorder; infection;
KW	neurological disease; drug screening; ss.
XX	
OS	Homo sapiens.
XX	
XX	WO200055350-A1.
PN	
XX	
PD	21-SEP-2000.
XX	
PP	08-MAR-2000; 2000WO-US055882.
PF	
XX	
XX	12-MAR-1999; 99US-0124270.
PR	
XX	
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
PI	Rosen CA, Ruben SM;
XX	
XX	WPI; 2000-587533/35.
XX	P-PSDB; AAB43971.
XX	
XX	Novel isolated nucleic acids comprising sequences encoding peptides
PT	useful for treating or diagnosing e.g. cancer -
PT	
PT	Claim 1: Page 1097-1098; 2352pp: English;
PS	
XX	AAC77607 to AAC78448 encode the human cancer associated proteins given
XX	in ABA43398 to ABA44239. The proteins can have activities based on the
CC	tissues and cells the genes are expressed in. Example of activities
CC	include: cytostatic; proliferative; vulnerable; immunomodulatory;
CC	antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC	antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC	dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC	neotropic; vasotropic; antipsoriatic and antiangiogenic. The
CC	polynucleotides and polypeptides can be used for preventing, treating or
CC	ameliorating medical conditions and diagnosing pathological conditions.
CC	Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC	the present invention may be used to treat immune disorders by activating
CC	or inhibiting the proliferation, differentiation or mobilisation of
CC	immune cells, to treat disorders of haematopoietic cells, autoimmune
CC	disorders, allergic reactions, graft versus host disease and organ
CC	rejection, modulate haemostatic or thrombolytic activity, modulate
CC	inflammation, cancers, cardiovascular disorders, neurological disease and
CC	bacterial or viral infections. The peptides, nucleotides, antibodies,
CC	agonists and antagonists may be also be used in drug screens. AAC78449 to
CC	AAC78457 and ABA44240 represent sequences used in the exemplification of
CC	the present invention.
XX	
XX	Sequence 2432 BP; 460 A; 764 C; 707 G; 496 T; 5 other;
SQ	

	Query Match	100.0%;	Score 627;	DB 21;	Length 2432;
	Best Local Similarity	100.0%;	Pred. No. 1,7e-154;	Mismatches 0;	Indels 0; Gaps 0;
	Matches 627;	Conservative	0;		
QY	1	gcaggcgacacctgtgagacccttgtagctgggtggccttaaacagacaggcgctggtggct	60		
Db	176	gcaggcgacacctgtgagacccttgtagctgggtggccttaaacagacaggcgctggtggct	235		
QY	61	ggaacgcggtgccagatggtcagttctgccttggctctgctggagaccgcggagga	120		
Db	236	ggaacgcggtgccagatggtcagttctgccttggctctgctggagaccgcggagga	295		
QY	121	gcagctacagctgctgccgtcccctcttgacaaaaggcccaacaactgagcgaggcat	180		
Db	296	gcagctacagctgctgccgtcccctcttgacaaaaggcccaacaactgagcgaggcat	355		
QY	181	ctgggtggccccctgccaggttgatgccactgctctgccggccaactcctgatctttacc	240		
Db	356	ctgggtggccccctgccaggttgatgccactgctctgccggccaactcctgatctttacc	415		
QY	241	gtctcagggaattccagtgtcgtccccttcocagagggccgtggagtcggggatggggcat	300		

Db	416	gtctcagggaacttcacagtgctgccccttcaccagaggccgtggcatgcggggatggcccat	478
QY	301	cactgtgtccacagggtctccactgcagtcagtcagacggcgcatctctgttccaaagatca	360
Db	476	cactgtgtccacagggtctccactgcagtcagtcagacggcgcatctctgttccaaagatca	535
QY	361	ggtacaaactcgtgggtgcccactccagtcgacctgatagtcagttcgaatgcccgacttc	420
Db	536	ggtaacaaactcgtgggtgcccactccagtcgacctgatagtcagttcgaatgcccgacttc	595
QY	421	tcacatgtgtgttatgtatgcattgctcctcctgggggtgctgcccacatgcccgacttc	480
Db	596	tcacatgtgtgttatgtatgcattgctcctcctgggggtgctgcccacatgcccgacttc	655
QY	481	tgctgtgaagacagggtgcactgctgcgcacggctgcctctgcgacctggttcacacc	540
Db	656	tgctgtgaagacagggtgcactgctgcgcacggctgcctctgcgacctggttcacacc	715
QY	541	cgctgcatcacacccacgggacaccccccctcctggcaagaagctcctgcccgaggact	600
Db	716	cgctgcatcacacccacgggacaccccccctcctggcaagaagctcctgcccgaggact	775
QY	601	aacagggcagtgctgttccagctcg	627
Db	776	aacagggcagtgctgttccagctcg	802
RESULT 5			
CC	AAQ14339	AAQ14339 standard; DNA; 1779 BP.	
XX	AAQ14339;		
XX	17-JAN-1992 (first entry)		
XX	Human epithelin precursor.		
XX	ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
XX	misc_RNA	41..1819	
XX	FT	/*tag= a	
XX	FT	/note= "claim 9, page 54"	
XX	PN	W09115510-A.	
XX	PD	17-OCT-1991.	
XX	PF	03-APR-1991; 91WO-US02321.	
XX	PR	13-MAR-1991; 91US-0083796.	
XX	PR	03-APR-1990; 90US-0504508.	
XX	PA	(BRIM ) BRISTOL-MYERS SQUIB.	
XX	PI	Shoyab M, Plowman GD;	
XX	XX	WPI; 1991-325168/44.	
XX	DR	P-PSDB; AARI4336.	
XX	PT	New cysteine-rich growth modulating proteins, epithelins - useful	
XX	PT	as inhibitors of neoplastic cell growth and to promote wound	
XX	PT	healing and treat psoriasis	
XX	PS	Disclosure; Fig 22; 97pp; English.	
XX	CC	ET-1 and ET-2 were isolated from rat kidneys and their amino acid	
XX	CC	sequences determined. A full length rat ET cDNA (AAQ14338) was obtnd.	
XX	CC	by screening a rat kidney cDNA library in lambda gt10 with PCR	
XX	CC	generated ET probes. These probes were also used to obtain the mouse	
XX	CC	ET gene (AAQ14340) from a mouse T-cell genomic library.	

ET DNA was also obtained from human sources (AAQ14339).  
An anti-sense ribonucleic acid molecule complementary to the  
indicated fragment in the features is also claimed.  
ET-1 is a bifunctional growth regulator, capable of stimulating  
the growth of some cell types while inhibiting the growth of others.  
ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
bioactivity. In contrast, however, ET-2 is apparently not capable of  
eliciting the growth stimulatory activity characteristic of ET-1 and,  
in fact, antagonises this ET-1 activity.  
See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.

```
Query Match      98.1%; Score 615; DB 12; Length 1779;
Best Local Similarity 100.0%; Pred. No. 2.1e-151;
Matches 615: Conservative 0; Mismatches 1 0; Indels 0; Gaps 0;
```

Qy	13	atgtggaacctggtgagctgggttgggaccttaacagcaggcgctggtggcttggaaacgcggtgc	72
Db	1	atgtggaacctggtgagctgggttgggaccttaacagcaggcgctggtggcttggaaacgcggtgc	60
Qy	73	ccagatgggtcaagttctgacctgtggcctgtgacctgaccccgaggagcagctacaagc	132
Db	61	ccagatgggtcaagttctgacctgtggcctgtgacctgaccccgaggagcagctacaagc	120
Qy	133	tgttcgcgtccctcttgacaataaggccacaacactgagcagcatctgggtggcccc	192
Db	121	tgttcgcgtccctcttgacaataaggccacaacactgagcagcatctgggtggcccc	180
Qy	193	tgcaggttgatgccactgctctgcggcactctgatatcttaccgctctcagggact	252
Db	181	tgcaggttgatgccactgctctgcggcactctgatatcttaccgctctcagggact	240
Qy	253	tccagtttgtcccccttcccagaggccgtggcaatggcgggatgggccatcaactgtctgccca	312
Db	241	tccagtttgtcccccttcccagaggccgtggcaatggcgggatgggccatcaactgtctgccca	300
Qy	313	cggggcttcacactgcaagtgcagacggcgatcctgcttccaaagatcaggttaacaactcc	372
Db	301	cggggcttcacactgcagtcagacggcgatcctgcttccaaagatcaggttaacaactcc	360
Qy	373	gtgggtgccatccagtgccctgatagtcagttcgaatgccgggaacttctccacgtgctgt	432
Db	361	gtgggtgccatccagtgccctgatagtcagttcgaatgccgggaacttctccacgtgctgt	420
Qy	433	gttatgtgatggctctctgggggtgctgcccatgccccagggttctctgctgtaagac	492
Db	421	gttatgtgatggctctctgggggtgctgcccatgccccagggttctctgctgtaagac	480
Qy	493	agggtgcacttgcttcgcgcaggtgcctcttcgcacctggttccacaccgcgtgcaccca	552
Db	481	agggtgcacttgcttcgcgcaggtgcctcttcgcacctggttccacaccgcgtgcaccca	540
Qy	553	cccacgggacccaccctctggcaagaagctcccttgcacaggactaacacaggcgatgc	612
Db	541	cccacgggacccaccctctggcaagaagctcccttgcacaggactaacacaggcgatgc	600

RESULT	6
AAQ14338	
ID	AAQ14338 standard; cDNA; 1767 BP.
XX	XX
XX	XX
AC	AAQ14338;
AC	XX
XX	XX
XX	17-JAN-1992 (first entry)
DT	XX
XX	XX
DE	Rat epithelin precursor.
XX	XX

ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.

Rattus rattus.

Key	Location/Qualifiers
misc_RNA	31...1797
	/ *tag= a
	/note= "claim 19, page 55"

WO9115510-A.

17-OCT-1991.

03-APR-1991; 91WO-US023321.

13-MAR-1991; 91US-0083796.

03-APR-1991; 90US-0504508.

(BRIM ) BRISTOL-MYERS SQUIB.

Shoyab M, Florman GD;

WPI: 1991-325168/44.

P-PSDB; AAR14325.

New cysteine-rich growth modulating proteins, epithelins - useful as inhibitors of neoplastic cell growth and to promote wound healing and treat psoriasis

Disclosure; Fig 18; 97pp; English.

ET-1 and ET-2 were isolated from rat kidneys and their amino acid sequences determined. A full length rat ET cDNA was obtd. by screening a rat kidney cDNA library in lambda gt10 with PCR generated ET probes. These probes were also used to obtain the mouse ET gene (AAQ14340) from a mouse R-cell genomic library. ET DNA was also obtained from human sources (AAQ14339)...

An anti-sense ribonucleic acid molecule complementary to the indicated fragment in the features is also claimed.

ET-1 is a bifunctional growth regulator, capable of stimulating the growth of some cell types while inhibiting the growth of others. ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory bioactivity. In contrast, however, ET-2 is apparently not capable of eliciting the growth stimulatory activity characteristic of ET-1 and, in fact, antagonises this ET-1 activity.

See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.

Sequence 1767 BP: 359 A: 499 C: 491 G: 418 T: 0 other;

Query Match	Score	DB 12;	Length
Best Local Similarity	64.8%	406.6;	1767;
Matches 488. Conservative	80.7%	Pred. No. 7.7e-97;	
		0: Mismatches 114;	Indels 3;
		Gaps 1;	

Qy	13	atgtggaccctggtgagctgggtgggcocttaacacagcaggcggtggtgctctggaacgcggtgc	72
Db	1	atgtgatcctggtgagctggctggccttagtgcaagctggtgctctggaacacagtcg	60
Qy	73	ccagatagtcagttctgcctctggtcctgctgctgagcccgagagccagccagctacagc	132
Db	61	ccagatggtcaattctgacctgttgcctgctgcttgaccaggaggagcccaactacagc	120
Qy	133	tgtcgcgtccctctctggacaaatggcccaacactgagcaggcatctgggtgggcccc	192
Db	121	tgtgtaacctctcttggaacatggcctataataacgagcgcgtctagtaggctcc	180
Qy	193	tgcaggttgatgccactctctgcgcgcacactctgcactcttaccgctcagggact	252
Db	181	tgccagatcccgtagccactgtctctgtagggtaactctgtctctcactgtgtctgggaact	240
Qy	253	tccagttgtgcctctccacagagccgtggcactgoggggatggccactcactgtctgccca	312
Db	241	tccactctaccctctctgaaggtgatctctgtgaatagggcagcactgtctgcc	300



AAQ14340 standard; DNA; 1767 BP.  
 AAQ14340;  
 17-JAN-1992 (first entry)  
 Mouse epithelin precursor.  
 ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.  
 Mus musculus.  
 Key Location/Qualifiers  
 8-1774  
 misc\_RNA /tag- a  
 /note= "claim 29, page 56"  
 WO9115510-A.  
 17-OCT-1991.  
 03-APR-1991; 91WO-US02321.  
 13-MAR-1991; 91US-0083796.  
 03-APR-1990; 90US-0504508.  
 (BRIM ) BRISTOL-MYERS SQUIB.  
 Shoyab M, Plowman GD;  
 WPI; 1991-325168/44.  
 P-PSDB; AAR14327.  
 New cysteine-rich growth modulating proteins, epithelins - useful  
 as inhibitors of neoplastic cell growth and to promote wound  
 healing and treat psoriasis  
 Disclosure; Fig 23; 97pp; English.  
 ET-1 and ET-2 were isolated from rat kidneys and their amino acid  
 sequences determined. A full length rat ET cDNA (AAQ14338) was obtd.  
 by screening a rat kidney cDNA library in lambda gt10 with PCR  
 generated ET probes. These probes were also used to obtain the mouse  
 ET gene (AAQ14340) from a mouse T-cell genomic library.  
 ET DNA was also obtained from human sources (AAQ14339).  
 An anti-sense ribonucleic acid molecule complementary to the  
 indicated fragment in the features is also claimed.  
 ET-1 is a bifunctional growth regulator, capable of stimulating  
 the growth of some cell types while inhibiting the growth of others.  
 ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
 bioactivity. In contrast, however, ET-2 is apparently not capable of  
 eliciting the growth stimulatory activity characteristic of ET-1 and,  
 in fact, antagonises this ET-1 activity.  
 See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
 Sequence 1767 BP; 360 A; 503 C; 492 G; 412 T; 0 other;  
 Query Match 62.3%; Score 390.4; DB 12; Length 1767;  
 Best Local Similarity 78.8%; Pred. No. 1.3e-92;  
 Matches 479; Conservative 0; Mismatches 126; Indels 3; Gaps 1;  
 Qy 13 atgtggacctggtgagctgggtggccttaacagcaggcgtgggtggcgaacgcggtgc 72  
 Db 1 atgtgggtcctgatgagctgggtggccttcgcaggcgtgggtggcgaacgcagctgt 60  
 Qy 73 ccagatggctactcctctgacctgctgacctgacctgacctgacctgacctgacctg 132  
 Db 61 ccagatggcagcttctgacctgctgacctgacctgacctgacctgacctgacctg 120  
 Qy 133 tgcctgcctcctctctggaacaaatggcccaacactgagcagcagcagcagcagc 192  
 Db 121 tgcgtaacctctctctggaacaaatggcccaacactgagcagcagcagcagcagc 180

Qy 193 tgccaggttgatcccaactgctctgcggccactcctgcactcttaccgtctcaggact 252  
 Db 181 tgccagaccatggccactgctcctgctgtattcttcttctcactgtgtctggact 240  
 Qy 253 tcagattgctgccccttccagagggcgtggcattgaggggagtgccatcactgctgccc 312  
 Db 241 tcagctgctgccccttctcctaagggtgtctgtgtgtggtggtaccactgctgccc 300  
 Qy 313 cggggcttccactgcagtcagacggcgagatcctgcttccaaagatcagatacaactcc 372  
 Db 301 cagggttccactgagtcagatgggaatcctgctccagatgtcagatacaactcc -- 358  
 Qy 373 gtgggtgctcactcagtcctgctgagtcagttcgaatgccggacttctcactgctgt 432  
 Db 359 -tggtgctgctcagtcctgctggagcagttgaaatgctgactgctgacactgctgc 417  
 Qy 433 gttatgctgagtcctgctgggtgctgcccattgcccaggtctcctgctgctggaagac 492  
 Db 418 attatggtgagttcgtggtgggtgtgtcccatgcccaggtcctgtgtggaagac 477  
 Qy 493 aggtgctgctgctgcagtcgcttctgacgtgcttgcacacccgctgcatcaca 552  
 Db 478 agagtgcattgctgctcccatggggcctcctgacgtgcttgcacacagatgcttca 537  
 Qy 553 cccacgggcccacccctggcgaagaagctccctgcccagagactaacagggcagtg 612  
 Db 538 cccacgggcccacccctactactaaagaagttcctgcacaaagacacacagggcagtg 597  
 Qy 613 gcttggc 620  
 Db 598 tctttgccc 605  
 RESULT 9  
 AAA69780/c  
 ID AAA69780 standard; cDNA; 561 BP.  
 AC AAA69780;  
 XX  
 XX 07-NOV-2000 (first entry)  
 XX Human ovarian carcinoma antigen polynucleotide SEQ ID NO:90.  
 DE Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;  
 KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.  
 KW Homo sapiens.  
 OS  
 XX WO200036107-A2.  
 XX 22-JUN-2000.  
 XX 17-DEC-1999; 99WO-US30270.  
 XX 17-DEC-1998; 98US-0215681.  
 XX 17-DEC-1998; 98US-0216003.  
 XX 23-JUN-1999; 99US-0338933.  
 XX 24-SEP-1999; 99US-0404879.  
 XX (CORI-) CORIXA CORP.  
 XX Mitcham JL, King GE, Algate PA, Frudakis TN;  
 XX WPI; 2000-431589/37.  
 XX Immuno-genic portion of an ovarian carcinoma protein and the nucleic  
 acid encoding it, useful for the diagnosis, prevention and treatment of  
 cancer, preferably ovarian cancer  
 XX Claim 18; Fig 15; 299pp; English.  
 XX The present invention describes an isolated polypeptide comprising an



KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX Homo sapiens.  
 XX EP1074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX (HELI-) HELIX RES INST.  
 PA Ota T, Isoqal T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs.  
 XX  
 PS Claim 8; SEQ ID 15309; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 1630 BP; 285 A; 517 C; 489 G; 339 T; 0 other;

Query Match 47.4%; Score 297.4; DB 22; Length 1630;  
 Best Local Similarity 99.7%; Pred. No. 2.9e-68;  
 Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 209 actgctgcggcgccactcctgcattcttaccgctcagggaacttccagttgctgccct 268  
 DB 1 actgctgcggcgccactcctgcattcttaccgctcagggaacttccagttgctgccct 60  
 QY 269 tccagagggccgtggcagtcgaggatgccatcactgctgccacgggggtccactgca 328  
 DB 61 tccagagggccgtggcagtcgaggatgccatcactgctgccacgggggtccactgca 120  
 QY 329 gtgcagagggccgtgcctgttccaaagatcaaggtaaacactcgtgggtgccatcaggt 388  
 DB 121 gtgcagagggccgtgcctgttccaaagatcaaggtaaacactcgtgggtgccatcaggt 180  
 QY 389 gccctgatgtcattcgaatgcccggaacttctccacgctgctgtgttatggctgagct 448

DB 181 gcctgatagtcagttcgaatgccggaacttctccacgctgctgttatgctgagct 240  
 QY 449 cctgggggtgctgccccatgccaggtctcctgctgagacaggtgctgctgctgct 507  
 DB 241 cctgggggtgctgccccatgccaggtctcctgctgagacaggtgctgctgctgct 299  
 RESULT 12  
 AAQ14953  
 ID AAQ14953 standard; DNA; 341 BP.  
 XX  
 AC AAQ14953;  
 XX  
 DT 17-JAN-1992 (first entry)  
 XX  
 DE Chicken epithelin precursor (partial).  
 XX  
 KW ET; growth regulation; inhibition; stimulation; ss.  
 XX  
 OS Gallus domesticus.  
 XX  
 PN W09115510-A.  
 XX  
 PD 17-OCT-1991.  
 XX  
 PF 03-APR-1991; 91WO-US02321.  
 XX  
 PR 13-MAR-1991; 91US-0083796.  
 PR 03-APR-1990; 90US-0504508.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIB.  
 XX  
 PI Shoyab M, Plowman GD;  
 XX  
 WPI; 1991-325168/44.  
 DR P-PSDB; AARI5427.  
 XX  
 XX New cysteine-rich growth modulating proteins, epithelins - useful  
 PT as inhibitors of neoplastic cell growth and to promote wound  
 PT healing and treat psoriasis  
 XX  
 PS Disclosure; Fig 25; 97pp; English.  
 XX  
 CC The epithelins appear to comprise several distinct members sharing  
 CC significant structural homology. Two members of the epithelin family,  
 CC EP-1 and EP-2, have been purified from natural sources, and cDNAs  
 CC encoding these and several other members of the epithelin family have  
 CC been isolated from rat (AAQ14338), human (AAQ14339), bovine (AAQ14952),  
 CC murine (AAQ14340) and chicken (AAQ14953).  
 CC EP-1 is a bifunctional growth regulator, capable of stimulating  
 CC the growth of some cell types while inhibiting the growth of others.  
 CC EP-2 is functionally similar to EP-1 w.r.t. growth inhibitory  
 CC bioactivity. In contrast, however, EP-2 is apparently not capable of  
 CC eliciting the growth stimulatory activity characteristic of EP-1 and,  
 CC in fact, antagonises this EP-1 activity.  
 CC See also AAQ14338-40, AAQ14952-53, AARI4328-9 and AARI5315-20.  
 XX  
 SQ Sequence 341 BP; 53 A; 108 C; 120 G; 59 T; 1 other;

Query Match 9.4%; Score 59; DB 12; Length 341;  
 Best Local Similarity 62.2%; Pred. No. 4.9e-06;  
 Matches 92; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 451 tgggggtgctgccccatgccaggtctcctgctgagacaggtgctgctgctgctgct 510  
 DB 4 tgggggtgctgccccatgccaggtgctgctgctgagacaggtgctgctgctgctgct 63  
 QY 511 caggtgcttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 570  
 DB 64 cactccaccagctgtgattggagcgcgctgctgctgctgctgctgctgctgctgctgct 123

Qy 571 ctggcaagaagctccctgcccagagga 598  
 ||||| || ||||| ||||| |||||  
 Db 124 atggcaccacaattccggcctggaaga 151

RESULT 13  
 AAC15693  
 ID AAC15693 standard; cDNA; 51 BP.  
 XX AC AAC15693;  
 XX DT 06-OCT-2000 (first entry)  
 XX DE Human secreted protein 5' EST, SEQ ID NO: 19768.  
 XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 XX KW gene therapy; chromosome mapping; ss.  
 XX OS Homo sapiens.  
 XX PN EP1033401-A2.  
 XX PD 06-SEP-2000.  
 XX PF 21-FEB-2000; 2000EP-0200610.  
 XX PR 26-FEB-1999; 99US-0122487.  
 XX PA (GEST ) GENSET.  
 XX PI Dunas Milne Edwards J, Duclert A, Giordano J;  
 XX DR WPI; 2000-500381/45.  
 XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX Claim 1; SEQ ID 19768; 7lpp + CD-ROM; English.  
 XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.  
 XX Sequence 51 BP; 12 A; 25 C; 8 G; 4 T; 2 other;

Query Match 8.0%; Score 50.2; DB 21; Length 51;  
 Best Local Similarity 96.1%; Pred. No. 0.0006;  
 Matches 49; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 537 caccgcgtgcatacaccacagggcaccacccctggcgaagaagctccc 587  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 1 caccgcgtgcatacaccacagggcaccacccctggcgaagaagctccc 51

RESULT 14  
 AAQ14952  
 ID AAQ14952 standard; DNA; 539 BP.  
 XX AC AAQ14952;  
 XX

DT 17-JAN-1992 (first entry)  
 XX Bovine epithelin precursor (partial).  
 XX DE  
 XX KW ET; growth regulation; inhibition; stimulation; ss.  
 XX OS Bos taurus.  
 XX PN WO9115510-A.  
 XX PD 17-OCT-1991.  
 XX PF 03-APR-1991; 91WO-US02321.  
 XX PR 13-MAR-1991; 91US-0083796.  
 XX PR 03-APR-1990; 90US-0504508.  
 XX PA (BRIM ) BRISTOL-MYERS SQUIB.  
 XX PI Shoyab M, Plowman GD;  
 XX DR WPI; 1991-325168/44.  
 XX DR P-PSDB; AAR15426.  
 XX PT New cysteine-rich growth modulating proteins, epithelins - useful  
 PT as inhibitors of neoplastic cell growth and to promote wound  
 PT healing and treat psoriasis  
 XX PS Disclosure; Fig 24; 97pp; English.  
 XX The epithelins appear to comprise several distinct members sharing  
 CC significant structural homology. Two members of the epithelin family,  
 CC EP-1 and EP-2, have been purified from natural sources, and cDNAs  
 CC encoding these and several other members of the epithelin family have  
 CC been isolated from rat (AAQ14338), human (AAQ14339), bovine, murine  
 CC (AAQ14340) and chicken (AAQ14953).  
 CC ET-1 is a bifunctional growth regulator, capable of stimulating  
 CC the growth of some cell types while inhibiting the growth of others.  
 CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
 CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
 CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
 CC in fact, antagonises this ET-1 activity.  
 CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
 XX Sequence 539 BP; 99 A; 171 C; 163 G; 106 T; 0 other;

Query Match 7.8%; Score 48.6; DB 12; Length 539;  
 Best Local Similarity 60.0%; Pred. No. 0.003;  
 Matches 81; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 394 gatagtcagttcgaatgcccgagcttctccacgtgctgtgtattgtgctggtcctcgg 453  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 169 gacatggaggtgagctgccagagcactacacctgtccgctcacagtcgggctgg 228

Qy 454 ggggtgctgcccattgcccgagcttctccacgtgctgtggaagacaggtgctgcgcac 513  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 229 ggcgtgctcctttgtgcagcgctgtgctgtgagaccattgtgcactgtgcgcgtcc 288

Qy 514 ggtgccttctgcac 528  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 289 gggtttaggtgtgcac 303

RESULT 15  
 AAS94337  
 ID AAS94337 standard; cDNA; 3729 BP.  
 XX AC AAS94337;  
 XX DT 13-FEB-2002 (first entry)  
 XX DE DNA encoding novel human diagnostic protein #30141.

Job time: 10533 sec

```
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG30150.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX Claim 1; SEQ ID No 30141; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS4197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 3729 BP; 691 A; 1150 C; 960 G; 928 T; 0 other;
```

Query Match 6.6%; Score 41.6; DB 23; Length 3729;  
Best Local Similarity 55.6%; Pred. No. 0.34;  
Matches 80; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

```
QY 206 cccactgctgcgcgcgcactctctgacatttaccgtctcaggagacttcagttgtctgcc 265
   || || || || || || || || || || || || || || || || || || || || ||
Db 3030 cctgtctctctgcgcgcctctctctgttttaacagcttctccacctctccacgagcctct 3089

QY 266 ccttcccagagcgcgtggtgcgcgggagtgccatcactgctccacgagggggttccact 325
   || || || || || || || || || || || || || || || || || || || || ||
Db 3090 cctgctcctggtgcgcctctctctgttttaacagcctctctctgttcaagtagcctctct 3149

QY 326 gcagtgcagacggtgcgcgtctctgct 349
   || || || || || || || || || || || || || || || || || || || || ||
Db 3150 cctgttcaacagcctctctctctct 3173
```

Search completed: September 5, 2002, 18:30:01



Result No.	Score	Query Match	Length	DB	ID	Description
1	627	100.0	2095	4	US-08-991-862-16	Sequence 16, Appl
2	615	98.1	1779	1	US-07-668-648-3	Sequence 3, Appl
3	615	98.1	1779	2	US-08-423-998-3	Sequence 3, Appl
4	615	98.1	1779	2	US-08-431-333-3	Sequence 3, Appl
5	615	98.1	1779	5	PT-0591-02321-3	Sequence 3, Appl
6	406.6	64.8	1767	1	US-07-668-648-1	Sequence 1, Appl
7	406.6	64.8	1767	2	US-08-423-998-1	Sequence 1, Appl
8	406.6	64.8	1767	2	US-08-431-333-1	Sequence 1, Appl
9	406.6	64.8	1767	5	PT-0591-02321-1	Sequence 1, Appl
10	399.2	63.7	2137	4	US-08-991-862-1	Sequence 1, Appl
11	390.4	62.3	1767	1	US-07-668-648-5	Sequence 5, Appl
12	390.4	62.3	1767	2	US-08-423-998-5	Sequence 5, Appl
13	390.4	62.3	1767	2	US-08-431-333-5	Sequence 5, Appl
14	390.4	62.3	1767	5	PT-0591-02321-5	Sequence 5, Appl
15	59	9.4	341	5	PT-0591-02321-9	Sequence 9, Appl
16	58.4	9.3	341	1	US-07-668-648-9	Sequence 9, Appl
17	58.4	9.3	341	2	US-08-423-998-9	Sequence 9, Appl
18	58.4	9.3	341	2	US-08-431-333-9	Sequence 9, Appl
19	48.6	7.8	539	1	US-07-668-648-7	Sequence 7, Appl
20	48.6	7.8	539	2	US-08-423-998-7	Sequence 7, Appl
21	48.6	7.8	539	2	US-08-431-333-7	Sequence 7, Appl
22	48.6	7.8	539	5	PT-0591-02321-7	Sequence 7, Appl
23	39.2	6.3	2580	3	US-09-050-863-2	Sequence 2, Appl
24	39.2	6.3	2580	4	US-09-359-081-2	Sequence 2, Appl
25	39.2	6.3	5452	2	US-09-130-114-1	Sequence 1, Appl
26	39.2	6.3	9600	4	US-08-910-647-1	Sequence 1, Appl
27	39.2	6.3	10596	1	US-07-884-841-15	Sequence 15, Appl

QY 301 cactgtgcccacggggttccactgcagtgacagcgggagatcctgcttccaaagatca 360  
Db 301 cactgtgcccacggggttccactgcagtgacagcgggagatcctgcttccaaagatca 360  
QY 361 ggtaaacactccgtggtgacatccagtgccctgagtcagttcgaaatgcccgacttc 420  
Db 361 ggtaaacactccgtggtgacatccagtgccctgagtcagttcgaaatgcccgacttc 420  
QY 421 tccacgtgtgttattggtgatggctcctctggtgggtgctgccccatgcccgaggttcc 480  
Db 421 tccacgtgtgttattggtgatggctcctctggtgggtgctgccccatgcccgaggttcc 480  
QY 481 tctgtgaagacaggggtgactgctgctccgacagtgcttctgacatggtttcacacc 540  
Db 481 tctgtgaagacaggggtgactgctgctccgacagtgcttctgacatggtttcacacc 540  
QY 541 cgtgcatcacaccacgggacccacccctggtgcaagaagctcctgcccagaggact 600  
Db 541 cgtgcatcacaccacgggacccacccctggtgcaagaagctcctgcccagaggact 600  
QY 601 aacaggcagtggttctgctcagctcg 627  
Db 601 aacaggcagtggttctgctcagctcg 627

## RESULT 2

US-07-668-648-3  
; Sequence 3, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/07668,648  
; FILING DATE: 19910819  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1779 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Kidney  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1779  
US-07-668-648-3

Query Match 98.1%; Score 615; DB 1; Length 1779;  
Best Local Similarity 100.0%; Pred. No. 1.3e-153;  
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 13 atgtgacacctgggtgagctgggtgagcttaacagagagggctggtgctggaacgcggtgc 72  
Db 1 ATGTGACACCTGGTGGCTGGGTGGCTTTAACAGACAGGGCTGGTGGCTGGAAACGGGTGC 60  
QY 73 ccagatgggtcagttctgcccctggtgctgctgacaccccgagagagcagctacagc 132  
Db 61 CCAGATGGGTCAAGTCTGCCCCCTGTGGCTGCTGCTGGAGACCCCGAGAGGCCACCTACAGC 120  
QY 133 tctgctgctccctctctggacaaatggcccaacacactgagcagggcatctgggtggtcccc 192  
Db 121 TGTGCTGCTCCCTTCTGGACAAATGGCCACAACTGAGCAGGCATCTGGGTGGGCCCC 180  
QY 193 tggcaggttgatgcccactgctgctgcccactcctgcatctttaccgtctcagggact 252  
Db 181 TGCACAGTTGATGTCGCCACTGCTGTGCCGCCACTCTCTGATCTTTACCGTCTCAGGGACT 240  
QY 253 tccagttgctgccccttcccagagcggcgtggcatgcgggatggtccatcactgctccca 312  
Db 241 TCCAGTTGCTGCCCCCTCCAGAGGCCGTGGCATGGGGATGGCCATCACTGCTGCCCA 300  
QY 313 cggggcttccactgcagtcagcgggctgctctctccaaagatcaggtaaacaactcc 372  
Db 301 CGGGGCTTCCACTGTCAGTCAGACGGCGGCGATCTCTCTCCAAAGATCAGGTAACTCC 360  
QY 373 gtgggtgccatccagtgcccctgatatgtagttcgaaatgcccgactctccacgtgctgt 432  
Db 361 GTGGGTGCCATCCAGTGCCTGATAGTCAGTTCGAATGCCCGGACTTCTCCAGGTGCTGT 420  
QY 433 gttatgtgcatggtctctgggggtgctgcccactcccacagcttctgctgtgaagac 492  
Db 421 GTTATGGTCATGGCTCTCTGGGGGTGCTGCCCATGCCCATGCTCTGCTGTGAAGAC 480  
QY 493 aggggtgactgctgctgcagcagtgctctctgacactggttcacaccccgctgcatcaaca 552  
Db 481 AGGGTGCATCTGCTGTCGGCAGCGTGTCTTCTGGACCTGGTTCACACCCCGCTGCATCACA 540  
QY 553 cccacgggcccaccccccctgccaagaagctccctgcccagaggactaacagggcagtg 612  
Db 541 CCCACGGGCACCCACCCCTGCGCAAGAAGCTCCCTGCCCGAGAGGACTTACAGGGCAGTG 600  
QY 613 gcctgtccagctcg 627  
Db 601 GCCTTGTCCAGCTCG 615

## RESULT 3

US-08-429-998-3  
; Sequence 3, Application US/08429998  
; Patent No. 585961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plowman, Gregory D.  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

121 TGCTGCCGTCCCTTCTGGACAAATGGCCCA

Db 541 CCCACGGGACCCACCCCCCTGGCAAGAAGCTCCCTGCCACAGGACTAACAGGGCAGTG 5000



```

Db 538 CCCACGGGCAACCCACCCTTACTTAAAGAAATTCCTCCCGCACAAAGGACCAAGGCAAGT 597
QY 613 gcttt 617
Db 598 GCTTT 602

RESULT 7
US-08-429-998-1
: Sequence 1, Application US/08429998
: Patent No. 5885961
: GENERAL INFORMATION:
: APPLICANT: Shoyab, Mohammed
: APPLICANT: Plowman, Gregory D.
: TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH
: TITLE OF INVENTION: MODULATING PROTEINS
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25.
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/429,998

```

```

? FILING DATE: 27-APR-1995
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/668,648
? FILING DATE: 13-MAR-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Misrock, S. Leslie
? REGISTRATION NUMBER: 18,872
? REFERENCE/DOCKET NUMBER: 5624-161-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-9741
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1767 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1767
? US-08-1998-1

? Query Match 64.8%; Score 406.6; DB 2; Length 1767;
? Best Local Similarity 80.7%; Pred. No. 1.4e-98;
? Matches 488; Conservative
?
? Qy 13 atgtggacctgtgagctgggtggccttaacagcaggcgtgtgctggaacgcggtgc 72
? ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
? Db 1 ATGTGGATCTCTGGTGACCTGGCTGGCCCTAGTGTGCGAAGCCTGCTGGCTGGACACAGTCG 60
?
? Qy 73 ccagatggctcagttctgccctgtggcctgtgctgtggaaccccgaggagccagctacagc 132
? ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
? Db 61 CCAGATGGTCAATTCTGCCCTGTGGCTGTGCTGTGCTTGTGACCAGGAGGAGCAACTACAG 120
?
? Qy 133 tgcgtgcgctccctcttcggacaataaggccacaacactgagcagcagcatctgagtggtgcccc 192
? ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
? Db 121 TCGTGTGAACCCCTCTTCCTGGACATGGCCTATTAATACAGAGCCGCTGCTTAGATGCTCC 180

```

Query Match	64.8%;	Score 405.6;	DB 2;	Length 1767;
Best Local Similarity	80.7%;	Pred. No. 1.4e-98;		
Matches 488: Conservative	0;	Mismatches 114;	Indels 3;	Gaps 1;



[illegible]

Query Match	63.7%	Score	399.2	DB	4	Length	2137
Best Local Similarity	78.9%	Pred. No.	1.4e-96				
Matches	489	Conservative	0	Mismatches	128	Indels	3
QY	1	cgcaggcagacatgtgagaccctgtgagctgtgagctgggtggccttaacacagcaggctgtggctt	60				
Db	11	cgcagacagacacatgtgggtctctgatgagctggtgcttcgcgcagcaggctggtagcc	70				
QY	61	ggaacgcgggtgccagatggtcagttctgcctgtggcctgtgctgtgcctggaccccgagga	120				
Db	71	ggaacacagtgccagatggcagttctgcctgttgctgtgctgtgacctgacaggagga	130				
QY	121	gccagctacagctgctgcctccctctctggacaaatggccacaacactggacaggcat	180				
Db	131	gccaaactacagctgctgtaacctctctggacacatggcctagaataacagagccatcat	190				
QY	181	ctgggtggccctgccaggttgatggccactgtctctgccggccactctgcatcttttacc	240				
Db	191	ctagatggctctgccagaccatggccactgtcctgtgctgtgctattctgtctctccat	250				
QY	241	gtctcaggagactccagttgctgccctctccagagccctggcatcggggatggccat	300				
Db	251	gtgtctggagactccagctgctgcctgtctctaaagggtgtgtctgtggtaggctac	310				
QY	301	cactgctgccacaggggtctccactgcagtgacagcgggcgagctgcttccaaagatca	360				
Db	311	cactgctccccagggtctccactgtagtgacatgggaaatcctgcttccagatgca	370				
QY	361	ggtaacaactcgtgggtgcaatccagtgccctgatagtcagttccgaatgcccgacttc	420				
Db	371	gataacct---tgggtgtgtccagtgctcctggggagccagttctgaatgctctgactct	427				
QY	421	tccacgtgctgtgttatgctcagtggtcctctgggggtgctgtgcccaatgccccaggcttc	480				

STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1767 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Rat  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1767  
PCT-US91-02321-1

Query Match	64.8%	Score	406.6	DB	5	Length	1767
Best Local Similarity	80.7%	Pred. No.	1.4e-98				
Matches	488	Conservative	0	Mismatches	114	Indels	3
Gaps	1						
QY	13	atgtggaacctgtggaagctgggtggcccttaacagcagggtgctgggtggatgggaacgcggtgc	72				
Db	1	ATGTGGATCTCTGTGAGCTGGCTGGCCCTTATGTGCCAAGGCTGTGTGGCTTGGAAACACATGTC	60				
QY	73	ccagatagctcagttctgcctctggtcctgctgctgcctggaccgccggagcagcagctacagc	132				
Db	61	CCAGATGGTCAATCTGCCTCTGTGCCTCTGTGCTTGCCTTGACCAGGGAGGAGGCCAATCTACAGC	120				
QY	133	tgtctgcgcgtccctctctgggacaaatgggcccaacaactgagcaggcatctgggtggcccc	192				
Db	121	TGCTGTAACCTCTCTCTGGACACATGGCTTATAATAACGAGCCGCTCGTCTAGATGGCTCC	180				
QY	193	tgcaggttgatgcccaactgctctgcggcccaactcctctgcactctttaccggtctcaaggact	252				
Db	181	TGCCAGATCCGTGACCACTGTCTGTATGGGTACTCTTGTCTCTCATTGTGTCTGGGACT	240				
QY	253	tccagttgctgccecttccagagccgtggcagtcgaggatggccatcactcgtgcgccca	312				
Db	241	TCCAGCTGTCTGCCGTCTCTGAGGGGTATCTTGTGTATGATGGCCAGCACTGCTGCCCC	300				
QY	313	cggggttccactgcagtcagacggcgatctctgtctccaaagatcaggtaacaactcc	372				
Db	301	CGGGGCTTCCACTCTAGTGGGATGGGAAATCTCTGCTCAGATATGATAGATAGCTCT	358				
QY	373	gtgggtgccatccagtgccttgatagtcagttcgaaatgcggagactctccagctgctgt	432				
Db	359	-TGGGTGCTGTCCAGTGTCTGTGTGAGCAGTTCGAATGTCTCTGACTCCGCCACCTGGTGT	417				
QY	433	gttatgtcagatgctcctgggggtgctgcctcccatgccccaggtctcgtctggaagac	492				
Db	418	ATTATGATGTATGGTCTCTGGGGGTCTTGCCCCATGCCCCAGGCCAGGGCTCTTGTCTGTGAAGAC	477				
QY	493	agggtgcactgctgcgcacggtgcctctctgcacactggttccacacccgctgatcaca	552				

RESULT 11  
US-07-668-648-5  
; Sequence 5, Application US/07686648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Plozman, Gregory D.  
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/07/668,648  
; APPLICATION NUMBER: US/07/668,648  
; FILING DATE: 19910819  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1767 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Mus musculus  
; TISSUE TYPE: Kidney  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1767  
; US-07-668-648-5

Query Match 62.3%; Score 390.4; DB 1; Length 1767;  
Best Local Similarity 78.8%; Pred No. 2.7e-94;  
Matches 479; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

Oy 13 atgtggacccttggtgagctgggtgccttaacagcaggctgggtcggaacgcggtgc 72  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 ATGTGGTCTCTGAGCTGCTGCTGCTTCGCCGCAGCGCTGTGCTAGCCGAACACAGTGT 60

RESULT 12  
US-08-429-998-5  
Sequence 5, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872

GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPTHELINS: NOVEL CYSTEINE-RICH GROWTH  
MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennile & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

;  
 ; FILING DATE: 27-APR-1995  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:

ATTACHMENT NUMBER: 03 07/000/040  
FILING DATE: 13-MAR-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090  
TELETYPE: (212) 960-0741

```

; IEEFAA: (212) 803-3741
; INFORMATION FOR SEQ ID NO: 5:

```

```

; SEQUENCE CHARACTERISTICS:
: LENGTH. 1767 base pairs

```

NAME: 107 2500 Faint  
; ;  
TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear

```

MOLECULE TYPE: cDNA

```

;
; ORIGINAL SOURCE:
; ORGANISM: MUS musculus
;

```

```

; TISSUE TYPE: Kidney
; FEATURE:

```

```

; ERROR.
; NAME/KEY: CDS

```

US-08-431-333-5  
; LOCATION: 1..1767

100

Query Match 62.3%; Score 390.4; DB 2; Length 1767;

Best Local Similarity	78.8%	Pred. No. 2.7e-94;	Index 3.
Matches	170.	Concentrative	126.
Matches	0.	Matches	126.

Matches 4/5/	Conservative	0;	Mismatches	120;	Index	3;	Outp

Qy 13 atgtggaccctggtgagctgggtggccttaacagcagggctggtggtggaacgcggtgc 72

Db 1 ATGTGGTCTGATGAGCTGGCTGGCCTTCGGCGCAGGGCTGGTAGCCGGACACAGTGT 60

Ov 73 ccagatggtcagtcttctgccctgtgcctqccctqccctqccctqccctacagc 132

120

**UD** CCAGATGGGCAGTTCTGCCCCGTTCCTGCTGCTGCTTGGACCCAGGGGAGGAGGCCCAATCATCAGC  
**bI**

Qy 133 tgctgccgtccccctctctggacaaaatggcccaactgagcaggcatctggtggcccc 192

Db 121 TGCTGTAACCCCTCTTCTGGACACATGGCCTAGAAATAACGAGCCATCATCTAGATGGCTCC 180

9v 193 taccaggttgatgcccactactctacgggccaactcctgcatctttaccgtctcagggaact 252

[illegible]

Db 181 TGCCAGACCCATGGCCACTGTCCCTGCTGGCTATTCTTGCTTCTCACACTGTGTGCTGGGACT 240

QY 253 tccagttgctgccccctccagagggccgtggcatgcgggagtgccatgcactgctgcca 312

Db 241 TCCAGCTGCTGCCCGTCTCTAAGGGTGCTCTTGTGGTGATGGCTACCACTGCTGCCCC 300

312 ~~see next page~~

372

[illegible]



ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Gallus domesticus  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..339  
PCT-US91-02321-9

Query Match 9.48; Score 59; DB 5; Length 341;  
Best Local Similarity 62.2%; Pred. No. 5.8e-07;  
Matches 92; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
Qy 451 tgggggtgctgccccatgccaggtctcctgctgtaaacacaggtgacactgctgccc 510  
||||| || ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||  
Db 4 TGGGGTTGTTGCCCGCCATGCCGNGAGCGGTGTGCTGCCGGGATGAGGAGCACTGTGTCCC 63  
Qy 511 cacgggtgcttctgacacctgttcacacccgctgcatcacacccagggcaccaccccc 570  
||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 64 CACTCCACCACCTGTGATTGGAGCGGGGGGCTGTGTCTCCCTACGGGGACGTCCCC 123  
Qy 571 ctggcaagaagctccctgcccagaga 598  
||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 124 ATGCCACCAAAATTCGGGGCCTGGAAGA 151

Search completed: September 5, 2002, 18:22:11  
Job time: 13334 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 13:08:22 ; Search time 3259.55 Seconds  
(without alignments)  
2596.245 Million cell updates/sec

Title: US-09-824-647-16\_COPY\_1\_627

Perfect score: 627

Sequence:

Scoring table: 1 cgcaggcagaccatgtggac.....cagtggcctgtccagctcg 627

IDENTITY\_NUC  
Gap 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627	100.0	653	10	BF725386
2	627	100.0	730	9	AU121637
3	627	100.0	760	9	AU130279
4	627	100.0	782	10	BI771292
5	627	100.0	785	10	BI771292
6	627	100.0	805	10	BM010556
7	627	100.0	814	9	AU130399
8	627	100.0	843	10	BI771009
9	627	100.0	849	9	AU131158
10	627	100.0	880	10	BI771009
11	627	100.0	920	9	AL552664
12	627	100.0	932	9	AL561424
13	627	100.0	986	9	AL548072
14	627	100.0	991	10	BM468251
15	627	100.0	1010	9	AL542702
16	627	100.0	1106	10	BM468512
17	627	100.0	1156	10	BM478783

18	625.4	99.7	763	10	BG824283
19	625.4	99.7	770	9	AU132413
20	625.4	99.7	782	10	BI755355
21	625.4	99.7	786	10	BG831133
22	625.4	99.7	804	9	AU125816
23	625.4	99.7	855	10	BM009189
24	625.4	99.7	882	10	BM007370
25	625.4	99.7	894	9	AL557765
26	624.4	99.6	874	10	BM051032
27	623.8	99.5	734	10	BG823225
28	623.8	99.5	770	10	BG469904
29	623.8	99.5	791	10	BI918228
30	623.8	99.5	958	10	BF312609
31	622.4	99.3	740	10	BM045357
32	622	99.2	789	9	AL525926
33	621.4	99.1	838	9	AU137266
34	620.2	98.9	677	9	AU121484
35	620.2	98.9	722	9	AU139905
36	620	98.9	788	9	AU124912
37	619	98.7	701	9	AU122545
38	618.8	98.7	805	9	AU141704
39	616.4	98.3	728	10	BE409050
40	616	98.2	700	10	BM046879
41	616	98.2	761	10	BG830081
42	616	98.2	834	9	AU122785
43	616	98.2	855	10	BM009145
44	616	98.2	1033	9	AL551350
45	615	98.1	774	10	BF312922

#### ALIGNMENTS

#### RESULT 1

BF725386

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF725386 653 bp mRNA linear EST 05-JAN-2001  
bx15c09.y1 Human Iris cDNA (Un-normalized, unamplified): BX Homo  
sapiens cDNA clone bx15c09 5', mRNA sequence.

BF725386.1 GI:12041297

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 653)

Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.

NEIBANK: EST analysis and bioinformatics for ocular genomics

Invest. Ophthalmol. Vis. Sci. 41 (2000) In press

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 15 row: c column: 09

Seq primer: M13Rpl reverse primer (ABI).

#### FEATURES

source

Location/Qualifiers

1..653

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="bx15c09"

BX

/tissue\_type="Iris"

/dev\_stage="Adult"

/lab\_host="EMDH10B"

/note="Organ: Eye; Vector: pcMVSPORT6; Post-mortem Iris

tissue was pooled from 10 individuals ranging in age from

4-80 years and RNA was extracted. From this pooled sample

an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A

directionally cloned cDNA library in the pcMVSPORT6 vector

Un-normalized, unamplified):

RESULT	2
AUI21637	
LOCUS	730 bp mRNA linear EST 19-OCT-2000
DEFINITION	MAMMAL Homo sapiens cDNA clone MAMMA1000615 5', mRNA sequence.
ACCESSION	AU121637
VERSION	AU121637.1 GI:10936872
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

	Matches	627;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	cgcaggcagaccat	tgtagaacct	ggtgagct	gggtgacctta	cagcagggct	ggtgacct	60		
Db	29	CGCAGGAGACAT	GTGGACCT	CGTGA	CTGGGTGGCTT	TAACAGCAGGGCT	GTGGTGGCT	88		
Qy	61	ggaacgcggtgcc	cagatggtc	agttctgc	ccctgtggc	ctgtgctg	ctgagacc	cgagga	120	
Db	89	GGAAACGGCGT	GCCAGATG	GTCTAGT	TTCGCCCT	GTGGCTGCT	GCCTGGAC	CCCGGAGGA	148	
Qy	121	gccagctacagct	gctgcctgc	ctcccttct	tgacaaat	tgcccaac	actgagcag	ggcat	180	
Db	149	GCCAGCTACAG	CTGCTGC	GTCCCTT	CTGGACAAAT	GGCCACAAC	ACTGAGCAGG	CAT	208	
Qy	181	ctgggtgcccct	gccagttgat	gccactgct	ctctgcgc	ccactctg	catctttacc	240		
Db	209	CTGGGTGCCCT	GCCAGGTT	GTATGCC	CACTGCTG	CCGCCAC	CTCTGTCAT	CTTTACC	268	
Qy	241	gtctcagggact	tcagttgct	gccccctt	ccccaggg	ccgtggcat	gcggggat	ggccat	300	
Db	269	GTCTCAGGGAC	TTCAGTTG	CTGCCCTT	TCCACAGAGGC	GTGGCATG	CGGGATGGCCAT	328		
Qy	301	cactgctgccac	ggggcttc	acactg	caatgcag	tcagacag	ggcgat	ctgcttcca	agatca	360
Db	329	CAC TGCTGCC	CACGGGGCTT	CCACTG	CACTGCAGTGC	ACAGGGCGG	CATCTGCTT	CTCAAGATCA	388	
Qy	361	ggtaacaact	ccgtggtg	gccatccag	tgccctgat	atgcatg	ttcgaat	ccccggacttc	420	
Db	389	GGTAACAAC	TCCGTGGTG	GCCATC	CAGTGCCCT	GATAGTCA	GTTCGAATG	CCCCGGACTTC	448	
Qy	421	tccacgtgctg	tgtatggt	gcgatg	gctctg	gggggtg	ctcccccat	gcccccaggtctcc	480	
Db	449	TCCACGTGCT	GTGTATGT	GTGGTGC	ATGGGTTC	CTTGGGGGT	GCTGCCCAT	TCCCGCAGGTTC	508	
Qy	481	tgctgtaaa	acaggg	tgactg	ctgcgc	acgggtg	cttctctg	cgactggtttcacacc	540	
Db	509	TGCTGTGA	AGACAGGG	TGCATG	CTGTCCG	CACGGTGC	CTTCTGGG	ACCTGGTTCACACC	568	
Qy	541	cgctgcatca	caccca	cgggc	accacca	ccccctt	ggcaaa	gaagctctctt	gccacagggact	600

Qy	301	cactgctgccacaggggcttccactgcaagtcagcagcggcgatcctctgcttccaaagatca	360		
Db	329	CACCTGCTGCCACAGGGGCTTCCACTGCAGTGCAGAGAGGGCGATCTCTTCCAAAGATCA	388		
Qy	361	ggtaaacactccctgggtggcattccactgcccctgattcagttcgaatgcccgagcttc	420		
Db	389	GGTAACAACATCCGCTGGGTGGCATCCAGTGCCTCTGATAGTCAGTTCGAATGCCCGGATTC	448		
Qy	421	tccacgtgctgtgttatgggtcgcattggtcctctgggggtgctgccccatgccccaggcttc	480		
Db	449	TCCACGTGCTGTGTATGTGTCGATGGCTCCTCTGGGGTGTCTGCCCATGCCCCAGGCTTC	508		
Qy	481	tgctgtgagacaggggtgcactgctgtccgcacggtgctctctgcagactgtgttcacacc	540		
Db	509	TGCTGTGAAGACAGGTCGCACTGCTGTGCCGACGGTGCCTTCTGCGACCTGTGTTCACCC	568		
Qy	541	cgctgatacacacccacgggcacccacccctggcgaagaagctccctgccagagact	600		
Db	569	CGCTGCAATCAACCCACGGGCACCCACCCCTGGCAAGAAGCTCCTGCCCAGGACT	628		
Qy	601	aacaggcactggcctgtgtccagctg	627		
Db	629	AACAGGCAGTGGCCTGTCTCAGCTCG	655		
RESULT 4					
LOCUS	BI771292	782 bp	mRNA linear EST 25-SEP-95		
DEFINITION	60304655F1 NIH_MGC_122	Homo sapiens	CDNA clone IMAGE:5204172		
ACCESSION	BI771292	mRNA sequence.			
VERSION	BI771292				
KEYWORDS	BI771292.1	GI:15762870			
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 782)				
TITLE	NIH-MGC http://www.ncbi.nlm.nih.gov/				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: qcapbs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1511 row: n column: 13 High quality sequence stop: 775. Location/Qualifiers 1. 782 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5204172" /clone_lib="NIH_MGC_122" /lab_host="DH10B"				
FEATURES	Source				
BASE COUNT	129 a	261 c	232 g 160 t		
ORIGIN	this is a NIH_MGC Library."				



```

ACCESSION      BM010556
VERSION        BM010556.1  GI:16524910
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
               National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-re@mail.nih.gov
               Tissue Procurement: DCTD/DTP
               cDNA Library Preparation: Ling Hong/Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLCM1933 row: 0 column: 11
               High quality sequence stop: 761.
FEATURES       Location/Qualifiers
               1..805
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:5448418"
               /clone_lib="NIH_MGC_41"
               /tissue_type="amelanotic melanoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
               EcoRI; cDNA made by oligo-dT priming. Directionally cloned
               into EcoRI/XhoI sites using the following 5' adaptor:
               GCCACGAG(G). Library constructed by Ling Hong in the
               laboratory of Gerald M. Rubin (University of California,
               Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
               Superscript II RT (Life technologies). Note: this is a
               NIH_MGC Library."
BASE COUNT     134 a 266 c 240 g 165 t
ORIGIN
Query Match    100.0%; Score 627; DB 10; Length 805;
Best Local Similarity 100.0%; Pred. No. 8.9e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  ccagcagaccatgacccctgtagctgagctgggtggtggttaacacagcaggctgtgct 60
Db 12  ccagcagaccatgtagcaccctgtagctgagctgggtggtggttaacacagcaggctgtgct 71
Qy 61  gaaacgcggtgcccagatggtcagttctgccctgtggtggtggtggtggtggtggtggtggt 120
Db 72  gaaacgcggtgcccagatggtcagttctgccctgtggtggtggtggtggtggtggtggtggt 131
Qy 121  gccagctacagctgtgcgtccctcttgcacaaatggccacacactgagcagggcat 180
Db 132  gccagctacagctgtgcgtccctcttgcacaaatggccacacactgagcagggcat 191
Qy 181  ctgggtgcccctgccagtgtagccacactgctcgtcgcgcgcacactcctgcatctttacc 240
Db 192  ctgggtgcccctgccagtgtagccacactgctcgtcgcgcgcacactcctgcatctttacc 251
Qy 241  gtctcaggagacttcagttgctgccccttcccagagggcgtggcagtcgqggatggccat 300
Db 252  gtctcaggagacttcagttgctgccccttcccagagggcgtggcagtcgqggatggccat 311
Qy 301  cactgctccccacggggttccactgcagtcgagcaggggcatcctgcttccaaagatca 360
Db 312  cactgctccccacggggttccactgcagtcgagcaggggcatcctgcttccaaagatca 371
Qy 361  gtaataactcgtgggtgcatccacagtcgacctgtagtcagttcgaatgcccgacttc 420
Db 361  gtaataactcgtgggtgcatccacagtcgacctgtagtcagttcgaatgcccgacttc 420

Db 372  GGTAACAACTCGGTGGTGCCTCAGTCCAGTGCCTGATGATGATGATGATGATGATGATGATGATGAT 431
Qy 421  tccagctgtgtgttatgttgatggtcctcctgggtggtggtggtggtggtggtggtggtggt 480
Db 432  tccagctgtgtgttatgttgatggtcctcctgggtggtggtggtggtggtggtggtggtggt 491
Qy 481  tgcgtgtaagacaggtgacgtgctgctgcacagctgcttctgcagctggtttcacacc 540
Db 492  tgcgtgtaagacaggtgacgtgctgctgcacagctgcttctgcagctggtttcacacc 551
Qy 541  cgtgcatcacaccacagggcagccacccctggcgaagaagctcctcgtgcccagaggact 600
Db 552  cgtgcatcacaccacagggcagccacccctggcgaagaagctcctcgtgcccagaggact 611
Qy 601  aacagggcagtggtcctgttccagctcg 627
Db 612  aacagggcagtggtcctgttccagctcg 638

RESULT 7
AUI30399
LOCUS       AUI30399 NT2RP3 Homo sapiens cDNA clone NT2RP3000775 5', mRNA
DEFINITION  AUI30399 NT2RP3 Homo sapiens cDNA clone NT2RP3000775 5', mRNA
ACCESSION   AUI30399
VERSION     AUI30399.1  GI:10990753
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 814)
AUTHORS     Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saiko,K., Kawal,Y.,
             Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
             Isogai,T.
TITLE       HRI human cDNA project
JOURNAL     Unpublished (2000)
COMMENT     Contact: Takao Isogai
             Genomics Laboratory
             Helix Research Institute
             1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
             Tel: 81-438-52-3951
             Fax: 81-438-52-3952
             Email: genomics@hri.co.jp
             HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix
             Research Institute; cDNA library construction: Department of
             Virology, Institute of Medical Science, University of Tokyo, and
             Helix Research Institute.
FEATURES    Location/Qualifiers
             1..814
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="NT2RP3000775"
             /clone_lib="NT2RP3"
             /cell_type="teratocarcinoma"
             /cell_line="NT2"
             /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
             cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT  139 a 267 c 237 g 166 t
ORIGIN
Query Match    100.0%; Score 627; DB 9; Length 814;
Best Local Similarity 100.0%; Pred. No. 8.9e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  ccagcagaccatgtagcaccctgtagctggtggtggtggtggtggtggtggtggtggtggt 60
Db 29  ccagcagaccatgtagcaccctgtagctggtggtggtggtggtggtggtggtggtggtggt 88
Qy 61  gaaacgcggtgcccagatggtcagttctgccctgtggtggtggtggtggtggtggtggtggt 120
Db 61  gaaacgcggtgcccagatggtcagttctgccctgtggtggtggtggtggtggtggtggtggt 120
Db 89  ggaacgcggtgcccagatggtcagttctgccctgtggtggtggtggtggtggtggtggtggt 148
```

```
121 gccagctacagctgctgcccctccctcttgagacaaatggcccaacactgacagcat 180
|||||
149 gccagctacagctgctgcccctccctcttgagacaaatggcccaacactgacagcat 208
|||||
181 ctgggtggccctgcccaggttgatgccactgtctgcccagcactctgcatcttaacc 240
|||||
209 CTGGGTGGCCCTGCCCAGGTGATGCCACTGCTGTGCCGGGCACCTCTCATCTTTTACC 268
|||||
241 gtctcaggagacttccagttgctcccttcccagagggcgtggcagtcgaggatggccat 300
|||||
269 GTCTCAGGAGACTTCCAGTTGCTGCCCTTCCAGAGGCCGTGCATGCGGGGATGGCCAT 328
|||||
301 cactgtgtcccaagggttccactgcagtcgagacgagggcgatctcttccaaagatca 360
|||||
329 CACTGTGTCCCGACGGGCTTCCACTGCAGTGCAGACAGGGCGATCTCTTCCAAAGATCA 388
|||||
361 ggttaacaaactccgtgggtgcatccagtcgctgagtcagttcgaatgcccgagcttc 420
|||||
389 GGTAAACAACTCCGTGGGTGCCATCCAGTGCCTGTATAGTCAGTTGGAATGCCGGAGCTTC 448
|||||
421 tccagctgctgtgtattggtcgatggctcctggtgggtgctgcccagtcgcccaggttcc 480
|||||
449 TCACAGTGTGTGTATTAGTGGCTGCTGTGGGGTGTGTGCCCATGCCCCAGGCTTCC 508
|||||
481 tgctgtgaagacaggggtgactgctgtccgcacagtgctcttctgacactggttcacacc 540
|||||
509 TGCTGTGAAGACAGAGGGTGCACCTGCTGTCCGCACGGTGCCTTCTGCGACCTTGGTTCAACC 568
|||||
541 cgctgtatcacacacacagggcaccacccctggcgaagaagctccctgcccagagact 600
|||||
569 CGTGTATCATCACACACAGGGACCCACCCCTGTGCCAAAGAGCTCTCTGCCACAGGACT 628
|||||
601 aacaggcgagtggtctgtccagctcg 627
|||||
629 AACAGGCGAGTGGCTGTCTCCAGCTCG 655

RESULT 8
BI771009 843 bp mRNA linear EST 25-SEP-2001
LOCUS 603055352F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204797 5',
DEFINITION mRNA sequence.
ACCESSION BI771009
VERSION BI771009.1 GI:15762587
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 843)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cyapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Cloned distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1513 row: h column: 14
High quality sequence stop: 843.
Location/Qualifiers
1..843
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5204797"
/lab_host="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
```

Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library.

BASE COUNT 134 a 283 c 252 g 174 t  
ORIGIN

```
Query Match 100.0%; Score 627; DB 10; Length 843;
Best Local Similarity 100.0%; Pred. No. 9e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgcagcagacacatgtagaccctggtagctgggtgagcttaacacagcagcgtggtgact 60
|||||
Db 64 CGCAGGCAGACCATGATGGACCCCTGGTGAGCTGGGTGGCTTTAACAGCAGGCTGGTGCT 123
|||||
QY 61 ggaacgcggtgcccagatggtcagttctgcccgtggtgctgctgagcccccggagga 120
|||||
Db 124 GGAACGGGTGCCAGATGTCAGTTCGCCCTGTGGCTGCTGCTGGACCCCGGAGGA 183
|||||
QY 121 gccagctacagctgctgcccgtcccccctcttggaacaaatggcccaacactgagcagcat 180
|||||
Db 184 GCCAGCTTACAGCTGCTGCCCTGCCCTTCTGTGACAAATGGCCCAACACTGAGCAGGCAT 243
|||||
QY 181 ctgggtggcccttgccaggtgatgccactgctctgcccggcactctgcatcttaacc 240
|||||
Db 244 CTGGGTGGCCCTTGCCAGGTGATGCCACTGCTCTGCCGGCCACTCTCTGCATCTTTTACC 303
|||||
QY 241 gtctcaggagacttccagttgctgccccttcccagagggcgtggcagtcgaggatggccat 300
|||||
Db 304 GTCTCAGGAGACTTCCAGTTGCTGCCCTTCCCAGAGGCCGTGGCATGCGGGGATGGCAT 363
|||||
QY 301 cactgtgtccacaggggttccactgcagtcgagacagggcgatcctgttccaaagatca 360
|||||
Db 364 CACTGTGTCCACAGGGGCTTCCACTGCAGTCGACAGCGGGCATCTGTCTTCCAAAGATCA 423
|||||
QY 361 ggttaacaaactccgtgggtgcatccactgcctgtagtcagttcgaatgcccgagcttc 420
|||||
Db 424 GGTAAACAACTCCGTGGGTGCCATCCAGTGCCTGTATAGTCAGTTGGAATGCCGGAGCTTC 483
|||||
QY 421 tccacgtgctgtgtattggtcgatggctcctggtgggtgctgcccagtcgcccaggtctcc 480
|||||
Db 484 TCCAGCTGCTGTATTAGTGGTGCATGGCTCTCTGGGGGTGCTGCCCCATGCCCGAGCTTCC 543
|||||
QY 481 tgctgtgaagacaggggtgactgctgtccgacaggtgctctctgacactggttccacacc 540
|||||
Db 544 TGCTGTGAAGACAGAGGTGCACCTGCTGTCCGCACAGGTGCTCTCTGCGACCTGTTGTTACACC 603
|||||
QY 541 cgctgtatcacacacacagggcaccacccctggcgaagaagctccctgcccagagact 600
|||||
Db 604 CGTGTATCATCACACACAGGGACCCACCCCTGTGCCAAAGAGCTCTCTGCCACAGGACT 663
|||||
QY 601 aacaggcgagtggtctgtccagctcg 627
|||||
Db 664 AACAGGCGAGTGGCTGTCTCCAGCTCG 690
|||||
```

```
RESULT 9
AUI31158
LOCUS AUI31158 NT2RP3 Homo sapiens cDNA clone NT2RP3002089 5', mRNA
DEFINITION sequence.
ACCESSION AUI31158
VERSION AUI31158.1 GI:10991512
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE 1 (bases 1 to 849)  
AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.  
TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3951  
Fax: 81-438-52-3952  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
ORGANISM Location/Qualifiers  
source  
1. 849  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="NT2RP3002089"  
/clone\_lib="NT2RP3"  
/cell\_type="teratocarcinoma"  
/cell\_line="NT2"  
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"  
BASE COUNT 143 a 279 c 248 g 175 t 4 others  
ORIGIN  
Query Match 100.0%; Score 627; DB 9; Length 849;  
Best Local Similarity 100.0%; Pred. No. 9e-135;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 cgcagcgagaccatgtgacccctggtgagctgggtggccttaacagcagggtgtggtc 60  
DB 29 ccgacgcagaccatgtgacccctggtgagctgggtggccttaacagcagggtgtggtc 88  
QY 61 ggaacgcggtgcccagatggtcagcttctgcttggcctggtggtcgtcctgaccccgagga 120  
DB 89 ggaacgcggtgcccagatggtcagcttctgcttggcctggtggtcgtcctgaccccgagga 148  
QY 121 gccagctacagctgctgcctgctcccttctggacaaatggccacacactgagcagcat 180  
DB 149 gccagctacagctgctgcctgctcccttctggacaaatggccacacactgagcagcat 208  
QY 181 ctgggtgcccctgcaggttatgcccactctctgcggccactcctgcatctttacc 240  
DB 209 ctgggtgcccctgcaggttatgcccactctctgcggccactcctgcatctttacc 268  
QY 241 gtctcagggaactccagttgtgcccctcccagagccgtggtggtggtggtggtggtc 300  
DB 269 gtctcagggaactccagttgtgcccctcccagagccgtggtggtggtggtggtggtc 328  
QY 301 cactgtgccacggggttccactgcagtcagtcagtcagtcagtcagtcagtcagtcagtc 360  
DB 329 cactgtgccacggggttccactgcagtcagtcagtcagtcagtcagtcagtcagtcagtc 388  
QY 361 ggtacaactccgtgggtggtggtggtggtggtggtggtggtggtggtggtggtggtc 420  
DB 389 ggtacaactccgtgggtggtggtggtggtggtggtggtggtggtggtggtggtggtc 448  
QY 421 tccagctgctgttatggtcagtcgctcgtggtggtggtggtggtggtggtggtggtc 480  
DB 449 tccagctgctgttatggtcagtcgctcgtggtggtggtggtggtggtggtggtggtc 508  
QY 481 tgcgtgaagacaggtgctgctgcgcagcagtcgctcgtggtggtggtggtggtggtc 540  
DB 509 tgcgtgaagacaggtgctgctgcgcagcagtcgctcgtggtggtggtggtggtggtc 568  
QY 541 cgtcgtacatacaacccacgggaccccccctcgtggcaagaagctccctgcccagaggt 600

DB 569 CCCTCATACACCCACGGGACCCACCCCTGGCAAGAGACTCCCTCCCGAGGACT 628  
QY 601 aacagggcagtggtccttgcagctcg 627  
DB 629 AACAGGGCAGTGGCCTGTGCCAGCTCG 655  
RESULT 10  
LOCUS BG823258  
DEFINITION 602726638F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4866201 5', mRNA sequence.  
ACCESSION BG823258  
VERSION BG823258.1 GI:14170845  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 880)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLC1732 row: 1 column: 10  
High quality sequence stop: 855.  
Location/Qualifiers  
source  
1. 880  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4866201"  
/clone\_lib="NIH\_MGC\_15"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pOTB7; Site:1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)".  
BASE COUNT 151 a 288 c 261 g 180 t  
ORIGIN  
Query Match 100.0%; Score 627; DB 10; Length 880;  
Best Local Similarity 100.0%; Pred. No. 9.2e-135;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 cgcagcgagaccatgtgacccctggtgagctgggtggtccttaacagcagggtggtggtc 60  
DB 12 ccgacgcagaccatgtgacccctggtgagctgggtggtccttaacagcagggtggtggtc 71  
QY 61 ggaacgcggtgcccagatggtcagttctgcccctggtggtcgtggtggtggtggtggtc 120  
DB 72 ggaacgcggtgcccagatggtcagttctgcccctggtggtcgtggtggtggtggtggtc 131  
QY 121 gccagctacagctgctgcctgctcccttctggacaaatggccacacactgagcagcat 180  
DB 132 gccagctacagctgctgcctgctcccttctggacaaatggccacacactgagcagcat 191  
QY 181 ctgggtgcccctgcaggttatgcccactctctgcggccactcctgcatctttacc 240



```

KEYWORDS      EST.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1. (bases 1 to 986)
AUTHORS        Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE          Full-length cDNA libraries and normalization
JOURNAL        Unpublished (2001)
COMMENT        Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 Evry cedex - France
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES       Location/Qualifiers
               1..986
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="CS0DI034YC17"
               /clone_lib="IT1.NFL006.PL2"
               /tissue_type="placenta"
               /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
               was primed with a NotI-oligo(dT) primer. Five prime end
               enriched, double-stranded cDNA was digested with Not I and
               cloned into the Not I and Eco RV sites of the pCMVSPORT 6
               vector. Library was normalized. Library was constructed by
               Life Technologies. Contact : Feng Liang Life Technologies,
               a division of Invitrogen 9800 Medical Center drive
               Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
               Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com"
BASE COUNT    168 a 323 c 290 g 203 t 2 others
ORIGIN
Query Match      100.0%; Score 627; DB 9; Length 986;
Best Local Similarity 100.0%; Pred. No. 9,5e-135;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cgcaggcagaccattgtggacccttggtgagctgggtggccttaacagcaggctggtggct 60
   |||||
Db 6 CGCAGGCAGACCATGTGGACCCCTGGTCAGCTGGGTGGCTCTTAACAGCAGGGCTGTGGCT 65
QY 61 ggaacggcgtgccagatgctcagtttgccttgccttgctgcctgcctgcctggaccccgaggga 120
   |||||
Db 66 GGAACGGCGTGCACAGATGTGAGTTGCGCCTGTGGCCTGTGCTGCTGCGTGACCCCGAGGA 125
QY 121 gccagctcacagctgctccgtccctctctggacaaatggccacaaacactgagcaggcat 180
   |||||
Db 126 GCACAGTACAGCTGCTGCCGTGCCCTTCTGGACAAATGGCCCAACACTGAGCAGGCAT 185
QY 181 ctgggtggcccttgccaggttgatgccactgctctgcggccactcgtcatctttacc 240
   |||||
Db 186 CTGGGTGCCCTTGCCAGGTTGATGCCACTGCTCTGCCGCCACTCCTGTGATCTTTACC 245
QY 241 gtctcagggaacttcagttgctgccttccagagccctggtgcatgcggggatggccat 300
   |||||
Db 246 GTCTCAGGGACTTCCAGTTGCTGCCCTTCCACAGAGGCCGTGGCATGCGGGATGGCCAT 305
QY 301 cactgctgccacggggtccactgcagtgcaagcggcgatcctcttccaaagatca 360
   |||||
Db 306 CACTGCTGCCACGGGGCTTCCACTGCAGTGCACAGGGCGCATCTGCTTCCAAGATCA 365
QY 361 ggtaaacactccgtgggtgccatccagtgccctgatagtcagttcgaatgccggacttc 420
   |||||
Db 366 GGTAAACAATCCGTGGGTGCCATCCAGTGCCTGTATAGTCAATTGATGCCGGACTTC 425
QY 421 tccacgtgctgtgttatggtcgatggctcctgggggtgctgcccccatgccccaggtctcc 480
   |||||
Db 426 TCCACGCTGCTGTATTATGGTCGATGGCTCCCTGGGGGTGCTGCCCATGCCCCAGGCTTCC 485
QY 481 tgcgtgaagacaggggtgactgctgtccgcaacgggtgacctctctgcagactggtttcacacc 540
   |||||
Db 486 TGCTGTGAACACAGGGGTGCATGTGTCGCCAGCGGTGCTCTTCGCACTCTGGTTACACACC 545

```

QY	241	gtctcagggaactccagttgctgccccctccagagagccgtgcatgctggggatggccat	300
Db	253	GTCTCAGGGGACTTCAGTTGCTGCCCCCTCCAGAGGCCGTGGCATGCGGGGATGGCCAT	312
QY	301	cactgctgccacaggggcttccactcagtgacagcagggcgatcctcttccaaagatca	360
Db	313	CAC TGCTGCCACAGGGGCTTCCACTGCAGTGACAGCGGGGATGCTCTCTCCAAAGATCA	372
QY	361	ggtaacaactccgtgggtggcattccactgagtcgctgatagtcagttcgaatccccggacttc	420
Db	373	GGTAACAACACTCCGTGGGTGCCATCCACTGCGCTGATAGTCAGTTCGAATGCCCGGACTTC	432
QY	421	tcaactgctgtgtattggtgatggtgctctgggggtgctgccccatgccccaggcttc	480
Db	433	TCCACGTGCTGTGTATTGTTGTCGATGGCTCTCTGGGGGTGCTGCCCATGCCCGAGCTTC	492
QY	481	tgctgtaaacagaggtgactctgtcccacaggttctctctgacacctgggttcacac	540
Db	493	TGCTGTGAAGACAGGGGTGCACTGCTGTGCGACGGTGCTCTCTGACACCTGGTTCACAC	552
QY	541	cgctgcatcacacccacgggcacccacccccctgtgcaagaagctccctgccccaggact	600
Db	553	CGCTGCATCACCCACGGGACCCACCCCTGCGAAAGAGCTCCCTGCCCGAGGACT	612
QY	601	aacagggcagtgccctgtgtccagctcg	627
Db	613	AACAGGGCAGTGGCCTTGTCCAGCTCG	639

RESULT 15

AL542702

LOCUS

DEFINITION

AL542702 LTI\_FL002\_PL1 Homo sapiens cDNA clone CS0E011YC02 5 prime

ACCESSION

AL542702

VERSION

AL542702.1

KEYWORDS

GI:12875005

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1010)

AUTHORS

L4, W. B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..1010

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0E011YC02"

/clone\_lib="LTI\_FL002\_PL1"

/lab\_host="DH10B"

/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT

170 a 327 c 300 g 211 t 2 others

ORIGIN

Query Match 100.0%; Score 627; DB 9; Length 1010;

Best Local Similarity 100.0%; Pred. No. 9.6e-135;



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 18:22:11 ; Search time 111.72 Seconds  
(without alignments)  
131.919 Million cell updates/sec

Title: US-09-824-647-16\_COPY\_796\_855

Perfect score: 60  
Sequence: 1 tccaaggagacgctaccac.....ctgcgcacacagtggcgcat 60

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2.6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	60	100.0	2095	4	US-08-991-862-16
2	58.4	97.3	1779	1	US-07-668-648-3
3	58.4	97.3	1779	2	US-08-429-998-3
4	58.4	97.3	1779	2	US-08-431-333-3
5	58.4	97.3	1779	5	PCT-US91-02321-3
6	34.4	57.3	539	1	US-07-668-648-7
7	34.4	57.3	539	2	US-08-429-998-7
8	34.4	57.3	539	2	US-08-431-333-7
9	34.4	57.3	539	5	PCT-US91-02321-7
10	31.2	52.0	1767	1	US-07-668-648-5
11	31.2	52.0	1767	2	US-08-429-998-5
12	31.2	52.0	1767	2	US-08-431-333-5
13	31.2	52.0	1767	5	PCT-US91-02321-5
14	31.2	52.0	2137	4	US-08-991-862-1
15	26.4	44.0	1767	1	US-07-668-648-1
16	26.4	44.0	1767	2	US-08-429-998-1
17	26.4	44.0	1767	2	US-08-431-333-1
18	26.4	44.0	1767	5	PCT-US91-02321-1
19	24.4	40.7	1377	4	US-09-377-557-17
20	23.4	39.0	4181	1	US-07-670-611-1
21	23.4	39.0	4181	1	US-08-220-674-1
22	23.4	39.0	4181	1	US-08-445-186-1
23	23.4	39.0	4181	1	US-08-446-549-1
24	23.4	39.0	4181	2	US-08-446-550-1
25	22.2	37.0	28958	1	US-08-258-261B-6
26	22.2	37.0	28958	1	US-08-456-837-6
27	22.2	37.0	28958	1	US-08-457-342-6

```

28 22.2 37.0 28958 1 US-08-457-646A-6 Sequence 6, Appli
29 22.2 37.0 28958 1 US-08-458-076A-6 Sequence 6, Appli
30 22.2 37.0 28958 1 US-08-764-233A-4 Sequence 4, Appli
31 22.2 37.0 28958 1 US-08-457-335A-6 Sequence 6, Appli
32 22.2 37.0 28958 1 US-08-729-214-6 Sequence 6, Appli
33 22.2 37.0 28958 3 US-09-028-934-6 Sequence 6, Appli
34 22.2 37.0 43377 1 US-08-764-233A-1 Sequence 1, Appli
35 22.2 37.0 441529 4 US-09-103-840A-1 Sequence 1, Appli
36 21.8 36.3 252 4 US-08-990-823-73 Sequence 73, Appli
37 21.8 36.3 1881 4 US-09-434-288-5 Sequence 5, Appli
38 21.8 36.3 7653 2 US-08-394-189B-1 Sequence 1, Appli
39 21.8 36.3 7653 3 US-08-258-287B-1 Sequence 1, Appli
40 21.8 36.3 7653 3 US-08-368-704C-1 Sequence 1, Appli
41 21.8 36.3 7653 5 PCT-US93-05701-18 Sequence 18, Appli
42 21.8 36.3 7653 5 PCT-US93-05705-1 Sequence 3, Appli
43 21.8 36.3 15664 1 US-08-402-282-3 Sequence 3, Appli
44 21.8 36.3 15664 1 US-08-508-004-3 Sequence 3, Appli
45 21.8 36.3 15664 1 US-08-402-066-3 Sequence 3, Appli

```

#### ALIGNMENTS

```

RESULT 1
US-08-991-862-16
; Sequence 16, Application US/08991862
; Patent No. 6309826
; GENERAL INFORMATION:
; APPLICANT: Serrero, GINETTE
; TITLE OF INVENTION: 88 KDA TUMORIGENIC GROWTH FACTOR AND ANTAGONISTS
; FILE REFERENCE: 29996.488/P001-A
; CURRENT APPLICATION NUMBER: US/08/991,862
; EARLIER FILING DATE: 1998-08-17
; EARLIER APPLICATION NUMBER: 08/863,862
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Human GP88 cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(1791)
; OTHER INFORMATION: Nucleotide sequence of human granulin/epithelin
; OTHER INFORMATION: precursor (human GP88). Human Granulin Genebank
; OTHER INFORMATION: M75161.
US-08-991-862-16

```

Query Match 100.0%; Score 60; DB 4; Length 2095;  
Best Local Similarity 100.0%; Pred. No. 2.3e-12;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 tccaaggagacgctaccacgacccctcctcactaagctgctgcgcacacagtggcgcat 60
|||||
Db 796 tccaaggagacgctaccacgacccctcctcactaagctgctgcgcacacagtggcgcat 855
|||||

```

```

RESULT 2
US-07-668-648-3
; Sequence 3, Application US/07668648
; Patent No. 5416192
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH
; TITLE OF INVENTION: MODULATING PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York

```

STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212) 869-9741  
TELEFAX: (212) 869-9741  
SEQUENCE CHARACTERISTICS:  
INFORMATION FOR SEQ ID NO: 3:  
LENGTH: 1779 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1779  
US-07-668-648-3

Query Match 97.3%; Score 58.4; DB 1; Length 1779;  
Best Local Similarity 98.3%; Pred. No. 8.2e-12;  
Matches 59; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaaggagaacgctaccacgacctctcactaagctgcctgcgcacacagtgaggcat 60  
|||||  
Db 784 TCCAAGGAGACGCTACCACGACCTCTCTACTAGCTGCCTGCCACACAGTGGGGGAT 843

RESULT 3  
US-08-429-998-3  
Sequence 3, Application US/08429998  
Patent No. 5885961  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429,998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212) 869-9741  
TELEFAX: (212) 869-9741  
SEQUENCE CHARACTERISTICS:  
INFORMATION FOR SEQ ID NO: 3:  
LENGTH: 1779 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1779  
US-08-429-998-3

Query Match 97.3%; Score 58.4; DB 2; Length 1779;  
Best Local Similarity 98.3%; Pred. No. 8.2e-12;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaaggagaacgctaccacgacctctcactaagctgcctgcgcacacagtgaggcat 60  
|||||  
Db 784 TCCAAGGAGACGCTACCACGACCTCTCTACTAGCTGCCTGCCACACAGTGGGGGAT 843

RESULT 4  
US-08-431-333-3  
Sequence 3, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212) 869-9741  
TELEFAX: (212) 869-9741  
SEQUENCE CHARACTERISTICS:  
INFORMATION FOR SEQ ID NO: 3:  
LENGTH: 1779 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; ORGANISM: Homo sapiens  
;; TISSUE TYPE: Kidney  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1779  
US-08-431-333-3

Query Match 97.3%; Score 58.4; DB 2; Length 1779;  
Best Local Similarity 98.3%; Pred. No. 8.2e-12;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tccaaggaagcgtaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60  
|||||  
Db 784 TCCAAGGAGAGCCTACCAAGGACCTCTCTACTAAGCTGCTGCCACACAGTGGGGAT 843

RESULT 5  
PCT-US91-02321-3  
; Sequence 3, Application PC/TUS9102321  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Bristol-Myers Squibb Company  
;; STREET: 3005 First Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98121

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US91/02321  
;; FILING DATE: 19910403  
;; CLASSIFICATION: 514

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Foot, Brian W.  
;; REGISTRATION NUMBER: 32,928  
;; REFERENCE/DOCKET NUMBER: ON0071A-PC  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206)728-4800  
;; TELEFAX: (206)448-4775

;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1779 base pairs  
;; TYPE: NUCLEIC ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; ORGANISM: Homo sapiens  
;; TISSUE TYPE: Kidney  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1779

PCT-US91-02321-3

Query Match 97.3%; Score 58.4; DB 5; Length 1779;  
Best Local Similarity 98.3%; Pred. No. 8.2e-12;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tccaaggaagcgtaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60

Db 784 TCCAAGGAGAGCCTACCAAGGACCTCTCTACTAAGCTGCTGCCACACAGTGGGGAT 843

RESULT 6  
US-07-668-648-7  
; Sequence 7, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/668,648  
;; FILING DATE: 19910819  
;; CLASSIFICATION: 514

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mistrock, S. Leslie  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 5624-161-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)790-9090  
;; TELEFAX: (212) 869-9741  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 539 base pairs  
;; TYPE: NUCLEIC ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; ORIGINAL SOURCE:  
;; ORGANISM: Bos taurus  
;; TISSUE TYPE: Kidney  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..537

US-07-668-648-7

Query Match 57.3%; Score 34.4; DB 1; Length 539;  
Best Local Similarity 85.0%; Pred. No. 0.0015;  
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Oy 1 tccaaggaagcgtaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60  
|||||  
Db 103 TCCAAGGAGAGCCT---ACGGACCTCTCTACCAAGCTCCCGCCGACACACAGTGCAGGAT 159

RESULT 7  
US-08-429-998-7  
; Sequence 7, Application US/08429998  
; Patent No. 5885961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/429-998  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Bos taurus  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..537  
US-08-429-998-7

Query Match 57.3%; Score 34.4; DB 2; Length 539;  
Best Local Similarity 85.0%; Pred. No. 0.0015;  
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
QY 1 tccaaggaagcgtaccacgacctctcactaagctgcgcacacagtggcgat 60  
|||||  
DB 103 TCCAGGAGAGCGCT---ACGGACCTCTCCACCAAGCTGCCCGCACACAGTGCAGGAT 159

RESULT 8  
US-08-431-333-7  
Sequence 7, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: PLOWMAN, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/431.333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Bos taurus  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..537  
US-08-431-333-7

Query Match 57.3%; Score 34.4; DB 2; Length 539;  
Best Local Similarity 85.0%; Pred. No. 0.0015;  
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 tccaaggaagcgtaccacgacctctcactaagctgcgcacacagtggcgat 60  
|||||  
DB 103 TCCAGGAGAGCGCT---ACGGACCTCTCCACCAAGCTGCCCGCACACAGTGCAGGAT 159

RESULT 9  
PCT-US91-02321-7  
Sequence 7, Application PC/TUS9102321  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: PLOWMAN, Gregory D.  
TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02321  
FILING DATE: 19910403  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Poor, Brian W.  
REGISTRATION NUMBER: 32,928  
REFERENCE/DOCKET NUMBER: ON0071A-PC  
TELEPHONE: (206)728-4800  
TELEFAX: (206)448-4775  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 base pairs

TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Bos taurus  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..537  
PCT-US91-02321-7

Query Match 57.3%; Score 34.4; DB 5; Length 539;  
Best Local Similarity 85.0%; Pred. No. 0.0015;  
Matches 51; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 1 tccaaggaacgtaccacgacctctactaagctgctgcgcacacagtgaggcgat 60  
Db 103 TCCAGGAGAACGCT---ACGACCTCTCACCACAGCTGCCCGCACACACAGTGCAGGAT 159

RESULT 10  
US-07-668-648-5  
; Sequence 5, Application US/07668648  
; Patent No. 5416192  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Flowman, Gregory D.  
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/668,648  
; FILING DATE: 19910819  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1767 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Mus musculus  
; TISSUE TYPE: Kidney  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1767  
US-07-668-648-5

Query Match 52.0%; Score 31.2; DB 1; Length 1767;  
Best Local Similarity 75.0%; Pred. No. 0.021;

Matches 39; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
Qy 8 agaagctaccacgacctctactaagctgctgcgcacacagtgaggcgat 59  
Db 785 AGAACTACACACGAGTCTCTGACCAAGCTGCTGGATACCCAGTGAAGGA 836

RESULT 11  
US-08-429-998-5  
; Sequence 5, Application US/08429998  
; Patent No. 5885961  
; GENERAL INFORMATION:  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Flowman, Gregory D.  
; TITLE OF INVENTION: EPIITHELINS: NOVEL CYSTEINE-RICH GROWTH  
; TITLE OF INVENTION: MODULATING PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/429,998  
; FILING DATE: 27-APR-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/668,648  
; FILING DATE: 13-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-161-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212) 869-9741  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1767 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Mus musculus  
; TISSUE TYPE: Kidney  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1767  
US-08-429-998-5

Query Match 52.0%; Score 31.2; DB 2; Length 1767;  
Best Local Similarity 75.0%; Pred. No. 0.021;  
Matches 39; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
Qy 8 agaagctaccacgacctctactaagctgctgcgcacacagtgaggcgat 59  
Db 785 AGAACTACACACGAGTCTCTGACCAAGCTGCTGGATACCCAGTGAAGGA 836

RESULT 12  
US-08-431-333-5  
; Sequence 5, Application US/08431333  
; Patent No. 5965723  
; GENERAL INFORMATION:



Db 807 agaactacacgagatctctcctgaccaagctgctggatataccacagtgaagga 858  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||

## RESULT 15

US-07-668-648-1

; Sequence 1, Application US/07668648

; Patent No. 5416192

; GENERAL INFORMATION:

; APPLICANT: Shoyab, Mohammed

; APPLICANT: Plozman, Gregory D.

; TITLE OF INVENTION: EPIRHELINS: NOVEL CYSTEINE-RICH GROWTH

; TITLE OF INVENTION: MODULATING PROTEINS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie &amp; Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/668,648

; FILING DATE: 19910819

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 5624-161-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)790-9090

; TELEFAX: (212) 869-9741

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1767 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1767

US-07-668-648-1

Query Match

44.0%; Score 26.4; DB 1; Length 1767;

Best Local Similarity

69.2%; Pred. No. 0.96;

Matches 36; Conservative

0; Mismatches 16;

Indels 0; Gaps

0;

Qy 8 agaacgtaccacgagctcctcactaagctgcctgcgcacacagtgggcga 59

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||

Db 785 AGGACTACACACAGATCTCATGACCAAGCTGCCCTGGATACCCAGCTGAATGA 836

Search completed: September 5, 2002, 18:22:14

Job time: 13337 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 18:30:01 ; Search time 457.59 seconds  
(without alignments)  
225.125 Million cell updates/sec

Title: US-09-824-647-16\_COPY\_796\_855

Perfect score: 60

Sequence: 1 tccaaggagaacgtaccac.....ctgcgcacacagtggcgat 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_032802.\*  
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	2095	AAV82825	Human GP88 autocri
2	60	100.0	2124	AAQ36794	Granulin coding se
3	60	100.0	2198	AAQ49052	Granulin coding se
4	58.4	97.3	757	AAH07141	Human cDNA clone (
5	58.4	97.3	1630	AAH16370	Human cDNA sequenc
6	58.4	97.3	1779	AAQ14339	Human epithelin pr
7	58.4	97.3	2432	AAQ78180	Human cancer assoc
8	34.4	57.3	539	AAQ14952	Bovine epithelin p
9	31.2	52.0	1767	AAQ14340	Mouse epithelin pr

10	31.2	52.0	2137	20	AAV82824	Mouse GP88 autocri
11	26.4	44.0	1767	12	AAQ14338	Rat epithelin prec
12	24.6	41.0	423	22	ABA44981	Human breast cell
13	24.6	41.0	423	22	ABA55457	Human foetal liver
14	24.6	41.0	423	22	ABA25175	Probe #3641 for ge
15	24.6	41.0	423	22	AAK03696	Human brain expres
16	24.6	41.0	423	22	AAK29156	Human bone marrow
17	24.6	41.0	423	22	AAI13744	Probe #3677 for ge
18	24.6	41.0	423	22	AAI35104	Probe #3790 used t
19	24.6	41.0	423	22	AAI03625	Probe #3616 used t
20	24.6	41.0	3512	20	AAV69395	H. contortus PGP-O
21	24.4	40.7	100	17	AAI30909	Primer F5 for 80 k
22	24.4	40.7	1377	22	AAI70511	Rice cryptophan de
23	24.4	40.7	2203	17	AAI30869	Engineered 80 kd p
24	23.8	39.7	1909	21	AAC49855	Arabidopsis thalia
25	23.4	39.0	159	22	ABA70391	Human foetal liver
26	23.4	39.0	159	22	AAK18628	Human brain expres
27	23.4	39.0	159	22	AAK4555	Human bone marrow
28	23.4	39.0	159	22	AAI50540	Probe #19226 used
29	23.4	39.0	441	23	AAI59469	DNA encoding novel
30	23.4	39.0	441	23	AAI59468	DNA encoding novel
31	23.4	39.0	488	22	ABA57774	Human foetal liver
32	23.4	39.0	488	22	AAK05839	Human brain expres
33	23.4	39.0	488	22	AAK31472	Human bone marrow
34	23.4	39.0	488	22	AAI37357	Probe #6043 used t
35	23.4	39.0	563	22	AAI505614	Mammalian vestibul
36	23.4	39.0	1550	21	AAI76777	Human ORFX ORF2332
37	23.4	39.0	1553	22	AAH13982	Human cDNA sequenc
38	23.4	39.0	1642	22	AAI08067	Human SPOF TRAF-pr
39	23.4	39.0	1642	22	AAH02922	Human shear stress
40	23.4	39.0	2012	22	AAI18136	Human DNA cytoxin
41	23.4	39.0	2141	23	AAI72373	DNA encoding novel
42	23.4	39.0	2789	21	AAI18162	Lung cancer associ
43	23.4	39.0	3978	23	ABL14929	Drosophila melanog
44	23.4	39.0	4181	13	AAQ29345	MCC gene of chromo
45	23.4	39.0	6557	23	ABL14928	Drosophila melanog

#### ALIGNMENTS

##### RESULT 1

AAV82825  
ID AAV82825 standard; cDNA; 2095 BP.

XX AAV82825;

XX 15-MAR-1999 (first entry)

XX Human GP88 autocrine growth factor cDNA.

XX GP88; granulin; epithelin; human; growth factor; autocrine; tumour;  
XX cancer; viral infection; antagonist; therapy; diagnosis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
XX CDS 13..1794  
XX /\*tag= a

XX WO9852607-AL.

XX 26-NOV-1998.

XX 22-MAY-1998; 98WO-US10555.

XX 16-DEC-1997; 97US-0991862.

XX 23-MAY-1997; 97US-0863079.

XX (SERR/) SERRERO G.

XX Serrero G;

DR WPI; 1999-045276/04.  
 DR P-PSDB; AAW85475.  
 XX Composition containing antagonist of growth factor GP88 - useful for  
 PT treating cancer and viral diseases and also for diagnosing disease  
 PT from altered GP88 expression  
 XX  
 XX Disclosure; Fig 9A; 86pp; English.  
 XX This human cDNA sequence includes a coding region for GP88 (see  
 CC AAW85474), an 88 kDa glycoprotein autocrine growth factor and  
 CC epithelin/granulin precursor that is expressed in a tightly  
 CC regulated manner in normal cells, is overexpressed and unregulated  
 CC in highly tumorigenic cells derived from normal cells, and which  
 CC acts as a stringently required growth stimulator for the  
 CC tumorigenic cells. Inhibition of GP88 expression or action in the  
 CC tumorigenic cells results in an inhibition of the tumorigenic  
 CC properties of the overproducing cells. Antagonists to GP88 are  
 CC used to treat diseases associated with increased expression of  
 CC GP88, particularly cancer but also viral infections. Fragments of  
 CC GP88 are used to raise specific antibodies (used as antagonists,  
 CC as diagnostic reagents and for delivering toxins or other  
 CC compounds to GP88-expressing cells) and to screen for antibodies.  
 CC Antisense oligonucleotides can also be used as antagonists.  
 CC Methods are provided for diagnosing disease, or determining  
 CC susceptibility to disease, resulting from altered GP88 activity.  
 XX  
 SQ Sequence 2095 BP; 364 A; 681 C; 625 G; 425 T; 0 other;

Query Match 100.0%; Score 60; DB 20; Length 2095;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 796 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 855

RESULT 2  
 AAQ56794  
 ID AAQ56794 standard; DNA; 2124 BP.  
 XX  
 AC AAQ56794;  
 XX  
 DT 22-APR-1994 (first entry)  
 XX  
 DE Granulin coding sequence.  
 XX  
 KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW granulocytes; leucocytes; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_difference 1379..1381  
 FT /\*tag= a  
 FT /transl\_except= ATG encodes Valine.  
 FT misc\_difference 1640..1642  
 FT /\*tag= b  
 FT /transl\_except= CAG encodes Glycine.  
 FT  
 XX WO9315195-A.  
 PN  
 XX  
 XX 05-AUG-1993.  
 PD  
 XX  
 XX 28-FEB-1992; 92WO-CA00089.  
 PF  
 XX  
 XX 03-FEB-1992; 92US-0829233.  
 PR  
 XX  
 XX (SOLO/) SOLOMON S.  
 PA  
 XX Solomon S;  
 PI  
 XX WPI; 1993-320328/40.  
 DR P-PSDB; AAR48673.  
 DR  
 XX New cystine rich granulin peptide(s) from leucocyte(s) - are  
 PT keratinocyte inhibitors useful topically for wound healing  
 PS Disclosure; Figure 4c; 53pp; English.  
 XX  
 XX The granulin inhibits keratinocytes and is useful in formulations  
 CC for promoting the healing of wounds.  
 CC  
 XX Sequence 2124 BP; 383 A; 685 C; 630 G; 426 T; 0 other;

XX WPI; 1993-320328/40.  
 DR P-PSDB; AAR48673.  
 XX  
 XX New cystine rich granulin peptide(s) from leucocyte(s) - are  
 PT keratinocyte inhibitors useful topically for wound healing  
 XX  
 XX Disclosure; Figure 4c; 53pp; English.  
 XX  
 XX The granulin inhibits keratinocytes and is useful in formulations  
 CC for promoting the healing of wounds.  
 CC  
 XX Sequence 2124 BP; 383 A; 685 C; 630 G; 426 T; 0 other;

Query Match 100.0%; Score 60; DB 14; Length 2124;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 809 tccaagagaaacgtaccacggacctctcactaagctgctgcgcacacagtggcgat 868

RESULT 3  
 AAQ49052  
 ID AAQ49052 standard; DNA; 2198 BP.  
 XX  
 AC AAQ49052;  
 XX  
 DT 22-APR-1994 (first entry)  
 XX  
 DE Granulin coding sequence.  
 XX  
 KW Granulin; keratinocytes; wound healing; inhibition; peptide;  
 KW granulocytes; leucocytes; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_difference 1453..1455  
 FT /\*tag= a  
 FT /transl\_except= ATG encodes Valine.  
 FT misc\_difference 1714..1716  
 FT /\*tag= a  
 FT /transl\_except= CAG encodes Glycine.  
 FT  
 XX WO9315195-A.  
 PN  
 XX  
 XX 05-AUG-1993.  
 PD  
 XX  
 XX 28-FEB-1992; 92WO-CA00089.  
 PF  
 XX  
 XX 03-FEB-1992; 92US-0829233.  
 PR  
 XX  
 XX (SOLO/) SOLOMON S.  
 PA  
 XX Solomon S;  
 PI  
 XX WPI; 1993-320328/40.  
 DR P-PSDB; AAR48673.  
 DR  
 XX New cystine rich granulin peptide(s) from leucocyte(s) - are  
 PT keratinocyte inhibitors useful topically for wound healing  
 PS Disclosure; Figure 4c; 53pp; English.  
 XX  
 XX The granulin inhibits keratinocytes and is useful in formulations  
 CC for promoting the healing of wounds.  
 CC  
 XX Sequence 2198 BP; 398 A; 708 C; 646 G; 446 T; 0 other;

Query Match 100.0%; Score 60; DB 14; Length 2198;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tccaaggaagcgtaccacgacctccctcactaagctgctgcgcacacagtgggcgat 60  
|||||  
Db 883 tccaaggaagcgtaccacgacctccctcactaagctgctgcgcacacagtgggcgat 942  
|||||

## RESULT 4

AAH07141  
ID AAH07141 standard; cDNA; 757 BP.

XX AAH07141;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:3976.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

PS Claim 1; SEQ ID 3976; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX Sequence 757 BP; 139 A; 232 C; 217 G; 165 T; 4 other;

Query Match 97.3%; Score 58.4; DB 22; Length 757;  
Best Local Similarity 98.3%; Pred. No. 7.2e-12;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggaagcgtaccacgacctccctcactaagctgctgcgcacacagtgggcgat 60  
|||||  
Db 328 tccaaggaagcgtaccacgacctccctcactaagctgctgcgcacacagtgggcgat 387  
|||||

## RESULT 5

AAH16370  
ID AAH16370 standard; cDNA; 1630 BP.

XX AAH16370;

XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:15309.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.  
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

PS Claim 8; SEQ ID 15309; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX Sequence 1630 BP; 285 A; 517 C; 489 G; 339 T; 0 other;  
SQ

Query Match 97.3%; Score 58.4; DB 22; Length 1630;  
Best Local Similarity 98.3%; Pred. No. 8.2e-12;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 60  
|||||  
Db 328 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 387  
|||||

RESULT 6  
AAQ14339  
ID AAQ14339 standard; DNA; 1779 BP.  
XX  
AC AAQ14339;  
XX  
DT 17-JAN-1992 (first entry)  
DE Human epithelin precursor.  
XX  
KW ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_RNA 41..1819  
/\*tag= a  
/note= "claim 9, page 54"  
XX  
XX WO9115510-A.  
XX  
XX 17-OCT-1991.  
XX  
XX 03-APR-1991; 91WO-US02321.  
XX  
XX 13-MAR-1991; 91US-0083796.  
XX  
XX 03-APR-1990; 90US-0504508.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIB.  
XX  
XX Shoyab M, Plozman GD;  
XX  
XX WPI; 1991-325168/44.  
XX  
XX P-PSDB; AAR14326.  
XX  
XX New cysteine-rich growth modulating proteins, epithelins - useful  
XX as inhibitors of neoplastic cell growth and to promote wound  
XX healing and treat psoriasis  
XX  
XX Disclosure; Fig 22; 97pp; English.  
XX  
XX ET-1 and ET-2 were isolated from rat kidneys and their amino acid  
XX sequences determined. A full length rat ET cDNA (AAQ14338) was obtd.  
XX by screening a rat kidney cDNA library in lambda gt10 with PCR  
XX generated ET probes. These probes were also used to obtain the mouse  
XX ET gene (AAQ14340) from a mouse T-cell genomic library.  
XX  
XX ET DNA was also obtained from human sources (AAQ14339).  
XX  
XX An anti-sense ribonucleic acid molecule complementary to the  
XX indicated fragment in the features is also claimed.  
XX  
XX ET-1 is a bifunctional growth regulator, capable of stimulating  
XX the growth of some cell types while inhibiting the growth of others.  
XX  
XX ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
XX bioactivity. In contrast, however, ET-2 is apparently not capable of  
XX eliciting the growth stimulatory activity characteristic of ET-1 and,  
XX in fact, antagonises this ET-1 activity.  
XX  
XX See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
XX  
XX Sequence 1779 BP; 304 A; 578 C; 547 G; 350 T; 0 other;  
SQ

Query Match 97.3%; Score 58.4; DB 12; Length 1779;  
Best Local Similarity 98.3%; Pred. No. 8.3e-12;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 60  
|||||  
Db 784 tccaagagaacgtaccacgacctctcactaagctgctgcgcacacagtgaggat 843  
|||||

RESULT 7  
AAC78180  
ID AAC78180 standard; CDNA; 2432 BP.  
XX  
AC AAC78180;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human cancer associated gene sequence SEQ ID NO:574.  
XX  
KW Human; cancer associated gene; cancer antigen; detection; cancer;  
XX diagnosis; cytostatic; proliferative; antineoplastic; immunomodulator;  
XX antidiabetic; antisthmatic; antirheumatic; antithrombotic; antiviral;  
XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;  
XX dermatological; neuroprotective; thrombolytic; coagulant; nootropic;  
XX vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;  
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;  
XX allergic reaction; graft versus host disease; organ rejection;  
XX haemostatic; thrombolytic; cardiovascular disorder; infection;  
XX neurological disease; drug screening; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200005350-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05882.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587533/55.  
XX  
XX P-PSDB; AAB43971.  
XX  
XX Novel isolated nucleic acids comprising sequences encoding peptides  
XX useful for treating or diagnosing e.g. cancer -  
XX  
XX Claim 1; Page 1097-1098; 2352pp; English.  
XX  
XX AAC78180 encodes the human cancer associated proteins given  
XX in AAB4398 to AAB44239. The proteins can have activities based on the  
XX tissues and cells the genes are expressed in. Example of activities  
XX include: cytostatic; proliferative; antineoplastic; immunomodulator;  
XX antidiabetic; antisthmatic; antirheumatic; antithrombotic;  
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
XX dermatological; neuroprotective; cardiac; thrombolytic; coagulant;  
XX nootropic; vasotropic; antipsoriatic and angiogenic. The  
XX polynucleotides and polypeptides can be used for preventing, treating or  
XX ameliorating medical conditions and diagnosing pathological conditions.  
XX  
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
XX the present invention may be used to treat immune disorders by activating  
XX or inhibiting the proliferation, differentiation or mobilisation of  
XX immune cells, to treat disorders of haematopoietic cells, autoimmune  
XX disorders, allergic reactions, graft versus host disease and organ  
XX rejection, modulate haemostatic or thrombolytic activity, modulate  
XX inflammation, cancers, cardiovascular disorders, neurological disease and  
XX bacterial or viral infections. The peptides, nucleotides, antibodies,  
XX agonists and antagonists may be also be used in drug screens. AAC78449 to  
XX AAC78457 and AAB44240 represent sequences used in the exemplification of  
XX the present invention.



ID AAV82824 standard; cDNA; 2137 BP.  
XX  
AC AAV82824;  
XX  
DT 15-MAR-1999 (first entry)  
XX  
DE Mouse GP88 autocrine growth factor cDNA.  
XX  
KW GP88; granulatin; epithelin; mouse; growth factor; autocrine; tumour;  
KW cancer; viral infection; antagonist; therapy; diagnosis; ss.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 23..1792  
FT FT /\*tag= a  
FT FT /transl\_except= (pos:44..45, aa:Ile)  
FT FT /transl\_except= (pos:182..184, aa:Asp)  
FT FT /transl\_except= (pos:1151..1153, aa:Cys)  
FT FT /transl\_except= (pos:1340..1342, aa:Leu)  
XX  
PN W09852607-A1.  
XX  
XX 26-NOV-1998.  
XX  
XX 22-MAY-1998; 98WO-US10555.  
XX  
PR 16-DEC-1997; 97US-0991862.  
PR 23-MAY-1997; 97US-0863079.  
XX  
XX (SERR/) SERRERO G.  
XX  
XX Serrero G;  
PI  
XX WPI; 1999-045276/04.  
DR P-PSDB; AAW85474.  
XX  
XX Composition containing antagonist of growth factor GP88 - useful for  
PT treating cancer and viral diseases and also for diagnosing disease  
PT from altered GP88 expression  
XX  
XX Example 5; Fig 8A-D; 86pp; English.  
XX  
XX This murine cDNA sequence includes a coding region for GP88 (see  
CC AAW85474), an 88 kDa glycoprotein autocrine growth factor and  
CC epithelin/granulin precursor that is expressed in a tightly  
CC regulated manner in normal cells, is overexpressed and unregulated  
CC in highly tumorigenic cells derived from normal cells, and which  
CC acts as a stringently required growth stimulator for the  
CC tumorigenic cells. Inhibition of GP88 expression or action in the  
CC tumorigenic cells results in an inhibition of the tumorigenic  
CC properties of the overproducing cells. Murine GP88 cDNA was  
CC isolated from the highly tumorigenic PC cell line cDNA library  
CC using a probe obtained by PCR using primers based on isolated GP88  
CC peptides. Antagonists to GP88 are used to treat diseases  
CC associated with increased expression of GP88, particularly cancer  
CC but also viral infections. Fragments of GP88 are used to raise  
CC specific antibodies (used as antagonists, as diagnostic reagents  
CC and for delivering toxins or other compounds to GP88-expressing  
CC cells) and to screen for antibodies. Antisense oligonucleotides  
CC can also be used as antagonists. Methods are provided for  
CC diagnosing disease, or determining susceptibility to disease,  
XX resulting from altered GP88 activity.  
XX  
XX Sequence 2137 BP; 443 A; 608 C; 583 G; 503 T; 0 other;  
SQ

Query Match 52.0%; Score 31.2; DB 20; Length 2137;  
Best Local Similarity 75.0%; Pred. NO. 0.066;  
Matches 39; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
Qy 8 agaagctaccacgagctctcactaagctgcctgcgcacacagtgagcgga 59  
||||| ||||||| ||||| || ||||||||| || ||||| ||

Db 807 agaactacacacgagatctctgaccaaagctgctggatataccacagtgagga 858  
RESULT 11  
AAQ14338  
ID AAQ14338 standard; cDNA; 1767 BP.  
XX  
XX AAQ14338;  
AC AAQ14338;  
XX  
XX 17-JAN-1992 (first entry)  
DT  
XX  
XX Rat epithelin precursor.  
DE  
XX  
XX ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.  
KW  
XX  
XX Rattus rattus.  
OS  
XX  
XX Key Location/Qualifiers  
FH misc\_RNA 31..1797 a  
FT FT /\*tag= a  
FT FT /note= "claim 19, page 55"  
XX  
XX W09115510-A.  
PN  
XX  
XX 17-OCT-1991.  
PD  
XX  
XX 03-APR-1991; 91WO-US02321.  
PF  
XX  
XX 13-MAR-1991; 91US-0083796.  
PR  
XX  
XX 03-APR-1990; 90US-0504508.  
PR  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIB.  
PA  
XX  
XX Shoyab M, Plowman GD;  
PI  
XX  
XX WPI; 1991-325168/44.  
DR  
XX  
XX P-PSDB; AAR14325.  
DR  
XX  
XX New cysteine-rich growth modulating proteins, epithelins - useful  
PT as inhibitors of neoplastic cell growth and to promote wound  
PT healing and treat psoriasis  
PT  
XX  
XX Disclosure; Fig 18; 97pp; English.  
PS  
XX  
XX ET-1 and ET-2 were isolated from rat kidneys and their amino acid  
CC sequences determined. A full length rat ET cDNA was obt'd. by screening  
CC a rat kidney cDNA library in lambda gt10 with PCR generated ET probes.  
CC These probes were also used to obtain the mouse ET gene (AAQ14340) from  
CC a mouse T-cell genomic library. ET DNA was also obtained from human  
CC sources (AAQ14339).  
CC An anti-sense ribonucleic acid molecule complementary to the  
CC indicated fragment in the features is also claimed.  
CC ET-1 is a bifunctional growth regulator, capable of stimulating  
CC the growth of some cell types while inhibiting the growth of others.  
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
CC in fact, antagonises this ET-1 activity.  
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
XX  
XX Sequence 1767 BP; 359 A; 499 C; 491 G; 418 T; 0 other;  
SQ

Query Match 44.0%; Score 26.4; DB 12; Length 1767;  
Best Local Similarity 69.2%; Pred. NO. 3.6;  
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
Qy 8 agaagctaccacgagctctcactaagctgcctgcgcacacagtgagcgga 59  
||||| ||||||| ||||| || ||||||||| || ||||| ||  
Db 785 aggaactacacacagatctcagacaaagctgcctggatataccacagtgagga 836  
||||| ||||||| ||||| || ||||||||| || ||||| ||

RESULT 12

RESULT 13  
ABA55457

```
OS Homo sapiens.
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 1; SEQ ID No 3641; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 423 BP; 110 A; 126 C; 100 G; 87 T; 0 other;

Query Match 41.0%; Score 24.6; DB 22; Length 423;
Best Local Similarity 70.2%; Pred. No. 13;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 5 aggagaacgtaccacgacctctcactaagctgcgtgcacaca 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 aggggaagataccaccagggttaacctcactaattgccctgcgcataca 222

RESULT 15
AAK03696
ID AAK03696 standard; DNA; 423 BP.
XX
XX AAK03696;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 3687.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
```

```
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 3687; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 423 BP; 110 A; 126 C; 100 G; 87 T; 0 other;

Query Match 41.0%; Score 24.6; DB 22; Length 423;
Best Local Similarity 70.2%; Pred. No. 13;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 5 aggagaacgtaccacgacctctcactaagctgcgtgcacaca 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 aggggaagataccaccagggttaacctcactaattgccctgcgcataca 222

Search completed: September 5, 2002, 18:30:02
Job time: 10534 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 15:34:05 ; Search time 3259.55 Seconds  
(without alignments)  
248.445 Million cell updates/sec

Title: us-09-824-647-16\_COPY\_796\_855

Perfect score: 60  
Sequence: 1 tccaaggagaactaccac.....ctgcgcacacagtggcgat 60

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_gss: \*  
13: em\_gss\_hum: \*  
14: em\_gss\_inv: \*  
15: em\_gss\_pln: \*  
16: em\_gss\_vrt: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	58.4	97.3	134	9 AW062972	AW062972 IL1-ST004
c 2	58.4	97.3	369	10 BF902443	BF902443 RC4-WT023
c 3	58.4	97.3	393	9 AA122084	AA122084 zm23e01.r
c 4	58.4	97.3	394	10 BF925544	BF925544 CM2-WT016
c 5	58.4	97.3	394	10 BG004140	BG004140 CM2-GN016
c 6	58.4	97.3	440	9 AI751894	AI751894 cn12e02.x
c 7	58.4	97.3	457	10 BF998533	BF998533 CM2-GN016
c 8	58.4	97.3	460	10 BG015966	BG015966 RC4-GN032
c 9	58.4	97.3	482	10 BF998626	BF998626 CM2-GN016
c 10	58.4	97.3	500	10 BG014975	BG014975 RC4-GN032
c 11	58.4	97.3	553	10 BE874824	BE874824 601488765
c 12	58.4	97.3	557	10 BF970308	BF970308 602273694
c 13	58.4	97.3	570	9 AW579074	AW579074 RC0-CT038
c 14	58.4	97.3	576	10 BI196113	BI196113 602754572
c 15	58.4	97.3	594	10 BI159891	BI159891 602863719
c 16	58.4	97.3	595	10 BE253430	BE253430 601111692
c 17	58.4	97.3	608	10 BE302516	BE302516 ba666003.y

```

18 58.4 97.3 630 10 BE741541
19 58.4 97.3 632 10 BI835123
20 58.4 97.3 641 10 BG423187
21 58.4 97.3 648 10 BM043523
22 58.4 97.3 657 10 BE742649
23 58.4 97.3 663 9 AU135509
24 58.4 97.3 664 10 BE743017
25 58.4 97.3 665 10 BE336895
26 58.4 97.3 677 10 BG748377
27 58.4 97.3 694 10 BE407219
28 58.4 97.3 696 10 BE258826
29 58.4 97.3 708 10 BI198717
30 58.4 97.3 716 10 BM048423
31 58.4 97.3 719 10 BI193612
32 58.4 97.3 737 10 BM043572
33 58.4 97.3 744 10 BF968260
34 58.4 97.3 757 9 AU134038
35 58.4 97.3 771 10 BI252407
36 58.4 97.3 776 10 BE386839
37 58.4 97.3 802 10 BE407885
38 58.4 97.3 803 10 BI870792
39 58.4 97.3 808 10 BG282501
40 58.4 97.3 813 10 BG753796
41 58.4 97.3 825 10 BF345137
42 58.4 97.3 838 10 BG284875
43 58.4 97.3 840 10 BM045888
44 58.4 97.3 844 10 BE383965
45 58.4 97.3 865 9 AL551474

```

#### ALIGNMENTS

```

RESULT 1
LOCUS AW062972/c
DEFINITION IL1-ST0041-020899-001-F07 ST0041 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW062972
VERSION AW062972.1 GI:6014357
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 134)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-IL1-ST0041-020
899-001-F07&t3=1999-08-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 134.
Location/Qualifiers
1. 134
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ST0041"
/dev_stage="Adult"
/note="*Organ: stomach; Vector: pUC18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters patent application
No. 196,716 - Ludwig Institute for Cancer Research)

```

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 27 a 39 c 34 g 34 t  
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 134;  
Best Local Similarity 98.3%; Pred. No. 3.8e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaaggagacgctaccacgacgtcctcctaagctgcctgcgcacacagtggcgat 60  
|||||  
Db 100 TCCAAGGAGACGCTACCGACGACCTCCTCACTAAGCTGCTCGGCACACAGTGGGGAT 41

RESULT 2  
BF902443/c  
LOCUS BF902443 369 bp mRNA linear EST 18-JAN-2001  
DEFINITION RC4-MT0235-131200-021-b04 MT0235 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF902443  
VERSION BF902443.1 GI:12293902  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 369)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4&t2=RC4-MT0235-  
131200-021-b04&t3=2000-12-13&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 52  
High quality sequence stop: 369.

FEATURES  
source  
1..369  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MT0235"  
/dev\_stage="Adult"  
/note="Organ: marrow; Vector: puc18; Site\_1: Smar; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 70 a 113 c 108 g 78 t  
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 369;  
Best Local Similarity 98.3%; Pred. No. 5.2e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagacgctaccacgacgtcctcctaagctgcctgcgcacacagtggcgat 60  
|||||  
Db 320 TCCAAGGAGACGCTACCGACGACCTCCTCACTAAGCTGCTCGGCACACAGTGGGGAT 261

RESULT 3  
AA122084  
LOCUS AA122084 393 bp mRNA linear EST 23-DEC-1997  
DEFINITION zM23e01.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone  
IMAGE:526488 5' similar to gb:X62320 GRANULINS PRECURSOR (HUMAN);,  
mRNA sequence.  
ACCESSION AA122084  
VERSION AA122084.1 GI:1678148  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 393)  
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins  
M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore  
B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,  
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Travaskis,E.,  
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
97044478

TITLE Contact: Wilson RK  
JOURNAL Washington University School of Medicine  
MEDLINE 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
COMMENT Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
WARNING: There is evidence that suggests that the 384-well parent  
plate of this clone contains both human and mouse derived clones.  
Thus, the origin of this clone is uncertain. This caution should be  
kept in mind should you use this clone.

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 1145 Std Error: 0.00  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 322.

FEATURES  
source  
1..393  
/organism="Homo sapiens"  
/db\_xref="GDB:3917917"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:526488"  
/clone\_lib="Stratagene pancreas (#937208)"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dr. Pancreatic adenocarcinoma cell line. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'  
CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT 68 a 118 c 117 g 89 t  
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 393;  
Best Local Similarity 98.3%; Pred. No. 5.3e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagacgctaccacgacgtcctcctaagctgcctgcgcacacagtggcgat 60  
|||||  
Db 176 TCCAAGGAGACGCTACCGACGACCTCCTCACTAAGCTGCTCGGCACACAGTGGGGAT 235

RESULT 4

BF925544 394 bp mRNA linear EST 19-JAN-2001  
 LOCUS CM2-NT0169-291100-525-c04 NT0169 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BF925544  
 ACCESSION BF925544  
 VERSION BF925544.1 GI:12321984  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

AUTHORS 1 (bases 1 to 394)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## MEDLINE

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-NT0169-  
 291100-525-c04&t3=2000-11-29&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 17

High quality sequence stop: 394.

Location/Qualifiers

1..394

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="NT0169"

/dev\_stage="Adult"

/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

71 a 123 c 118 g 82 t

## BASE COUNT

## ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 394;  
 Best Local Similarity 98.3%; Pred. No. 5.3e-09;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tcacaggagacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60  
 |||||

Db 205 TCCAGGAGACGCTACACGACGACCTCTCTACTAAGCTGCTGCGCACACAGTGGGGAT 264

RESULT 5

LOCUS BG004140

DEFINITION CM2-GN0166-201100-525-c04 GN0166 Homo sapiens cDNA, mRNA sequence.

ACCESSION BG004140

VERSION BG004140.1 GI:12445017

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 394)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-NT0169-  
 291100-525-c04&t3=2000-11-29&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 17

High quality sequence stop: 394.

Location/Qualifiers

1..394

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="NT0169"

/dev\_stage="Adult"

/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

71 a 123 c 118 g 82 t

Query Match 97.3%; Score 58.4; DB 10; Length 394;  
 Best Local Similarity 98.3%; Pred. No. 5.3e-09;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tcacaggagacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60  
 |||||

Db 205 TCCAGGAGACGCTACACGACGACCTCTCTACTAAGCTGCTGCGCACACAGTGGGGAT 264

RESULT 6

LOCUS AT751894

DEFINITION clone NT2e02.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA

clone NT2e02 random, mRNA sequence.

ACCESSION AT751894

VERSION AT751894.1 GI:5130158

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 440)

Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G.,  
 Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey  
 , P.G., Hotchkiss, R.N. and Francomano, C.A.  
 SGAP: The Skeletal Genome Anatomy Project  
 Unpublished (1997)  
 Contact: Libin Jia  
 Medical Genetics Branch  
 National Human Genome Research Institute

## REFERENCE

## AUTHORS

1 (bases 1 to 394)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663

## JOURNAL

## MEDLINE

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-GN0166-  
 291100-525-c04&t3=2000-11-20&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 17

High quality sequence stop: 394.

Location/Qualifiers

1..394

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="GN0166"

/dev\_stage="Adult"

/note="Organ: placenta\_normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent  
 application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the pUC 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."

71 a 123 c 118 g 82 t

Query Match 97.3%; Score 58.4; DB 10; Length 394;  
 Best Local Similarity 98.3%; Pred. No. 5.3e-09;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tcacaggagacgtaccacgacctctcactaagctgcctgcgcacacagtggcgat 60  
 |||||

Db 205 TCCAGGAGACGCTACACGACGACCTCTCTACTAAGCTGCTGCGCACACAGTGGGGAT 264

RESULT 6

LOCUS AT751894

DEFINITION clone NT2e02.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA

clone NT2e02 random, mRNA sequence.

ACCESSION AT751894

VERSION AT751894.1 GI:5130158

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 440)

Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G.,  
 Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey  
 , P.G., Hotchkiss, R.N. and Francomano, C.A.  
 SGAP: The Skeletal Genome Anatomy Project  
 Unpublished (1997)  
 Contact: Libin Jia  
 Medical Genetics Branch  
 National Human Genome Research Institute

10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA  
 Tel: 301-402-4877  
 Fax: 301-496-7157  
 Email: libhelix.nih.gov  
 DNA Sequencing and analyses by National Institutes of Health  
 Intramural Sequencing Center (NISC).  
 Plate: 12 row: e column: 02  
 Seq primer: -21M13 forward primer (ABI).  
 Location/Qualifiers

## FEATURES

source

1..440  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="NHTBC.cml2e02"  
 /clone\_lib="Normal Human Trabecular Bone Cells"  
 /sex="Female"  
 /tissue\_type="Bone"  
 /cell\_type="Trabecular Bone Cells"  
 /lab\_host="SURE"  
 /note="Organ: Hip; Vector: pBluescript; Site\_1: EcoRI"  
 BASE COUNT 81 a 149 c 125 g 85 t  
 ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 440;  
 Best Local Similarity 98.3%; Pred. No. 5.5e-09;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 tccaaggagaacgtaccacgacctctactaagctgctgcgcacacagtggcgat 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 298 TCCAAGGAGACGCTACACGACCTCTCTACTAAGCTGCTGCACACAGTGGGGAT 357

## RESULT 7

LOCUS

BF998533  
 CM2-GN0166-101100-496-f05 GN0166 Homo sapiens cDNA, mRNA sequence.  
 BF998533  
 BF998533.1 GI:12434060  
 EST.  
 KEYWORDS  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 457)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2at2-CM2-GN0166-101100-496-f05&t3=2000-11-10&t4=1>)

Seq primer: puc 18 forward  
 High quality sequence start: 24  
 High quality sequence stop: 328.

Location/Qualifiers  
 1..457  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

## FEATURES

source

1..457  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

/clone\_lib="GN0166"  
 /dev\_stage="Adult"

/note="Organ: placenta\_normal; Vector: puc18; Site\_1: SmaI  
 ; Site\_2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent  
 application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the pUC 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."  
 BASE COUNT 97 a 139 c 135 g 86 t  
 ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 457;  
 Best Local Similarity 98.3%; Pred. No. 5.6e-09;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tccaaggagaacgtaccacgacctctactaagctgctgcgcacacagtggcgat 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 281 TCCAAGGAGACGCTACACGACCTCTCTACTAAGCTGCTGCACACAGTGGGGAT 222

## RESULT 8

LOCUS

BG015966  
 RC4-GN0323-281200-032-g01 GN0323 Homo sapiens cDNA, mRNA sequence.  
 BG015966  
 BG015966.1 GI:12468699  
 EST.  
 KEYWORDS  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 460)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663

COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4st2-RC4-GN0323-281200-032-g01&t3=2000-12-28&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence start: 103  
 High quality sequence stop: 460.

## FEATURES

source

1..460  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="GN0323"  
 /dev\_stage="Adult"

/note="Organ: placenta\_normal; Vector: puc18; Site\_1: SmaI  
 ; Site\_2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent  
 application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the pUC 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."  
 BASE COUNT 85 a 135 c 149 g 91 t

## ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 460;  
 Best Local Similarity 98.3%; Pred. No. 5.7e-09;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagctaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60  
 |||||  
 DB 181 TCCAGGAGAAGCTACCACGAGCTCCTCACTAGCTGCCTGGCGCACACAGTGGGGGAT 240

## RESULT 9

BF998626 482 bp mRNA linear EST 24-JAN-2001  
 LOCUS CM2-GN0166-101100-496-g05 GN0166 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BF998626  
 ACCESSION BF998626  
 VERSION BF998626.1 GI:12434248  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 482)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil

Tel: +55-11-2704922  
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-GN0166-  
 101100-496-g05&t3=2000-11-10&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 20

High quality sequence stop: 130.

Location/Qualifiers

## FEATURES

source

1..482  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="GN0166"  
 /dev\_stage="Adult"

/note="Organ: placenta\_normal; Vector: puc18; Site\_1: SmaI  
 ; Site\_2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent  
 application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the puc 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."

BASE COUNT 94 a 142 c 146 g 100 t

## ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 482;  
 Best Local Similarity 98.3%; Pred. No. 5.7e-09;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagctaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60  
 |||||  
 DB 221 TCCAGGAGAAGCTACCACGAGCTCCTCACTAGCTGCCTGGCGCACACAGTGGGGGAT 280

## RESULT 10

BG014975

LOCUS RC4-GN0323-211200-021-f01

DEFINITION GN0323 Homo sapiens cDNA, mRNA sequence.

ACCESSION BG014975

VERSION BG014975.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 500)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC4&t2=RC4-GN0323-  
 211200-021-f01&t3=2000-12-21&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 35

High quality sequence stop: 500.

Location/Qualifiers

1..500

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="GN0323"

/dev\_stage="Adult"

/note="Organ: placenta\_normal; Vector: puc18; Site\_1: SmaI

; Site\_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the puc 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

BASE COUNT 103 a 160 c 139 g 98 t

ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 500;

Best Local Similarity 98.3%; Pred. No. 5.7e-09;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagctaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60

|||||

DB 140 TCCAGGAGAAGCTACCACGAGCTCCTCACTAGCTGCCTGGCGCACACAGTGGGGGAT 199

|||||

RESULT 11

BE874824

LOCUS 601488765F1 NTH\_MSC\_69

DEFINITION Homo sapiens cDNA clone IMAGE:3890969 5',

KEYWORDS mRNA sequence.

ACCESSION BE874824

VERSION BE874824.1

KEYWORDS EST. GI:10323600

BASE COUNT 553 bp mRNA linear EST 20-OCT-2000

Query Match 97.3%; Score 58.4; DB 10; Length 500;

Best Local Similarity 98.3%; Pred. No. 5.7e-09;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagctaccacgacctctcactaagctgcctgcgcacacagtgggcgat 60

|||||

DB 140 TCCAGGAGAAGCTACCACGAGCTCCTCACTAGCTGCCTGGCGCACACAGTGGGGGAT 199

|||||

SOURCE  
ORGANISM

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 553)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: DCTD/DTG/Gazdar  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM9675 row: a column: 18  
High quality sequence stop: 540.

Location/Qualifiers  
1. .553  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3890969"  
/clone\_lib="NIH\_MGC\_69"  
/tissue\_type="large cell carcinoma, undifferentiated"  
/lab\_host="DH10B (phage-resistant)"  
/notes="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."  
102 a 178 c 162 g 111 t

FEATURES  
source

Query Match 97.3%; Score 58.4; DB 10; Length 553;  
Best Local Similarity 98.3%; Pred. No. 5.9e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaagagacgtaccacggacctctcactaagctgctgcgcacagtgggcgat 60  
|||||  
Db 173 TCCAAGGAGAACGCTACCGGACCTCTCACTAAGCTGCTGCGCACAGTGGGGGAT 232

BASE COUNT  
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 553;  
Best Local Similarity 98.3%; Pred. No. 5.9e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaagagacgtaccacggacctctcactaagctgctgcgcacagtgggcgat 60  
|||||  
Db 173 TCCAAGGAGAACGCTACCGGACCTCTCACTAAGCTGCTGCGCACAGTGGGGGAT 232

RESULT 12  
BF970308

LOCUS 557 bp mRNA linear EST 22-JAN-2001  
DEFINITION 602273694F1 NIH\_MGC\_84 Homo sapiens CDNA clone IMAGE:4361638 5',  
mRNA sequence.  
ACCESSION BF970308  
VERSION BF970308.1 GI:12337596  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 557)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10004 row: 1 column: 23  
High quality sequence stop: 556.

FEATURES  
source

Query Match 97.3%; Score 58.4; DB 9; Length 570;  
Best Local Similarity 98.3%; Pred. No. 6e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaagagacgtaccacggacctctcactaagctgctgcgcacagtgggcgat 60  
|||||  
Db 458 TCCAAGGAGAACGCTACCGGACCTCTCACTAAGCTGCTGCGCACAGTGGGGGAT 517

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4361638"  
/clone\_lib="NIH\_MGC\_84"  
/tissue\_type="adrenal cortex carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/notes="Organ: adrenal gland; Vector: pCMV-SPORT6; Site\_1:  
NotI; Site\_2: SalI; Cloned unidirectionally. Oligo-dT  
primed. Average insert size 1.229 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
105 a 182 c 155 g 115 t

BASE COUNT  
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 557;  
Best Local Similarity 98.3%; Pred. No. 5.9e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaagagacgtaccacggacctctcactaagctgctgcgcacagtgggcgat 60  
|||||  
Db 458 TCCAAGGAGAACGCTACCGGACCTCTCACTAAGCTGCTGCGCACAGTGGGGGAT 517

RESULT 13  
AW579074

LOCUS 570 bp mRNA linear EST 16-MAR-2000  
DEFINITION RC0-CT0380-070100-031-g04 CT0380 Homo sapiens CDNA, mRNA sequence.  
ACCESSION AW579074  
VERSION AW579074.1 GI:7254123  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 570)  
HCGP <http://www.ludwig.org.br/ORESTES>.  
The FAPESP/LICR Human Cancer Genome Project  
Unpublished (1999)  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0&t2=RC0-CT0380-070100-031-g04&t3=2000-01-07&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 278  
High quality sequence stop: 570.

FEATURES  
source

Location/Qualifiers  
1. .570  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0380"  
/dev\_stage="Adult"  
/notes="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
109 a 183 c 159 g 119 t

BASE COUNT  
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 570;  
Best Local Similarity 98.3%; Pred. No. 6e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaagagacgtaccacggacctctcactaagctgctgcgcacagtgggcgat 60  
|||||  
Db 458 TCCAAGGAGAACGCTACCGGACCTCTCACTAAGCTGCTGCGCACAGTGGGGGAT 517

FEATURES  
source

Location/Qualifiers  
1. .570  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0380"  
/dev\_stage="Adult"  
/notes="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
109 a 183 c 159 g 119 t

BASE COUNT  
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 570;  
Best Local Similarity 98.3%; Pred. No. 6e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaagagacgtaccacggacctctcactaagctgctgcgcacagtgggcgat 60  
|||||  
Db 458 TCCAAGGAGAACGCTACCGGACCTCTCACTAAGCTGCTGCGCACAGTGGGGGAT 517

FEATURES  
source

Location/Qualifiers  
1. .557  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0380"  
/dev\_stage="Adult"  
/notes="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
109 a 183 c 159 g 119 t

BASE COUNT  
ORIGIN

Query Match 97.3%; Score 58.4; DB 9; Length 570;  
Best Local Similarity 98.3%; Pred. No. 6e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 tccaagagacgtaccacggacctctcactaagctgctgcgcacagtgggcgat 60  
|||||  
Db 458 TCCAAGGAGAACGCTACCGGACCTCTCACTAAGCTGCTGCGCACAGTGGGGGAT 517

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagcgtaccacggacctctcactaaagctgctgcgcacacagtgggcgat 60  
|||||  
Db 468 TCCAAGGAGAACGCTACCACGGACCTCCTCACTAAGCTGCTGCGCACACAGTGGGGGAT 527

RESULT 14  
BI196113  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BI196113  
602754572F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4890472 5',  
mRNA sequence.  
BI196113  
BI196113.1 GI:14651133  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 576)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
Plate: L16M1767 row: 0 column: 17  
High quality sequence stop: 575.

FEATURES  
Location/Qualifiers  
1..576  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4890472"  
/clone\_lib="NIH\_MGC\_19"  
/tissue\_type="neuroblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
ECORI; CDNA made by oligo-dT priming. Directionally  
cloned into EORI/XhoI sites using the following 5',  
adaptor: GGCACGAG(C). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 107 a 189 c 164 g 116 t  
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 576;  
Best Local Similarity 98.3%; Pred. No. 6e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagcgtaccacggacctctcactaaagctgctgcgcacacagtgggcgat 60  
|||||  
Db 487 TCCAAGGAGAACGCTACCACGGACCTCCTCACTAAGCTGCTGCGCACACAGTGGGGGAT 546

RESULT 15  
BI159891  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BI159891  
602863719F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:5017733 5',  
mRNA sequence.  
BI159891  
BI159891.1 GI:14619892  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 594)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L16M1830 row: f column: 06  
High quality sequence start: 6  
High quality sequence stop: 582.

FEATURES  
Location/Qualifiers  
1..594  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5017733"  
/clone\_lib="NIH\_MGC\_42"  
/tissue\_type="epithelioid carcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EORI; CDNA made by oligo-dT priming.  
Directionally cloned into EORI/XhoI sites using the  
following 5' adaptor: GGCACGAG(C). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 119 a 182 c 181 g 112 t  
ORIGIN

Query Match 97.3%; Score 58.4; DB 10; Length 594;  
Best Local Similarity 98.3%; Pred. No. 6e-09;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tccaaggagaagcgtaccacggacctctcactaaagctgctgcgcacacagtgggcgat 60  
|||||  
Db 95 TCCAAGGAGAACGCTACCACGGACCTCCTCACTAAGCTGCTGCGCACACAGTGGGGGAT 154

Search completed: September 5, 2002, 15:34:08  
Job time: 8746 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 15:34:08 ; Search time 3259.55 Seconds  
(without alignments)  
4438.875 Million cell updates/sec

Title: US-09-824-647-16\_COPY\_1024\_2095

Perfect score: 1072

Sequence: 1 gggccaccaggcgccctg.....ataaagttgtcactttctt 1072

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	918	85.6	1018	9	AL572883
C 2	904	84.3	996	9	AL582634
C 3	860.8	80.3	942	9	AL542659
C 4	817.8	76.3	900	9	AL574944
C 5	765	71.4	887	9	AL576991
C 6	760.4	70.9	810	10	BG684182
C 7	757	70.6	908	9	AL562633
C 8	747	69.7	808	9	AL579730
C 9	734.8	68.5	835	9	AL575030
C 10	713	66.5	780	10	BG763617
C 11	695	64.8	745	10	BG281390
C 12	693.4	64.7	807	10	BE742164
C 13	681.2	63.5	894	10	BE7424853
C 14	673.2	62.8	751	10	BG830299
C 15	667.6	62.3	888	10	BE253057
C 16	664.6	62.0	877	10	BG284114
C 17	662.2	61.8	806	10	BG742228

18	658.4	61.4	745	10	BI562752
19	649	60.5	706	10	BE743031
20	647.4	60.4	878	10	BG765435
21	639.8	59.7	732	10	BE747981
22	639.2	59.6	1082	10	BM466432
C 23	636.8	59.4	753	9	AL570288
C 24	611.4	57.0	714	10	BM007936
C 25	604.6	55.4	753	9	AW131849
C 26	596.8	56.7	687	10	BI831698
C 27	596.6	55.7	734	9	AI985672
C 28	593.6	55.4	967	10	BG110722
C 29	593.4	55.4	766	10	BI200080
C 30	584.8	54.6	769	10	BI252854
C 31	584.6	54.5	628	9	BE042869
C 32	581.8	54.3	920	10	BG770670
C 33	579.6	54.1	710	10	BG251638
C 34	575.6	53.7	702	10	BF981909
C 35	575.2	53.7	639	9	BE042879
C 36	573.4	53.5	925	9	AI375908
C 37	568	53.0	881	9	AL541797
C 38	567.2	52.9	652	9	AW263746
C 39	565.8	52.8	715	9	AI887551
C 40	564.4	52.6	727	10	BE272013
C 41	563.6	52.6	714	10	BG032688
C 42	561.2	52.4	877	10	BE899344
C 43	556.8	51.9	599	10	BG818217
C 44	549.6	51.3	649	9	AW166916
C 45	549.2	51.2	666	9	AI888176

#### ALIGNMENTS

RESULT 1  
AL572883/c  
LOCUS  
DEFINITION AL572883 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0DI034YC17 3 prime, mRNA sequence.  
AL572883  
VERSION AL572883.1 GI:12931579  
KEYWORDS EST.  
SOURCE Human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1018)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
JOURNAL  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
Location/Qualifiers  
1..1018  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DI034YC17"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"  
BASE COUNT 199 a 306 c 331 g 172 t 10 others  
ORIGIN

BI562752 603235401  
BE743031 601574660  
BG765435 602738917  
BE747981 601580466  
BM466432 AGENCOURT  
AL570288 AL570288  
BM007936 603617658  
AW131849 xf35d08.x  
BI831698 603078851  
AI985672 wt19f02.x  
BG110722 602279017  
BI200080 602760780  
BI252854 602953251  
BE042869 ho30a08.x  
BG770670 602734010  
BG251638 602363526  
BF981909 602308625  
BE042879 ho30b07.x  
AI375908 tc14c02.x  
AL541797 AL541797  
AW263746 xq47g12.x  
AI887551 wml5a12.x  
BE272013 601141195  
BG032688 602298331  
BE899344 601681181  
BG818217 602779818  
AW166916 xg68d01.x  
AI888176 wml4b06.x



638 GTGTGTCGAGGATGCCAGCACTGCTGCCCGGTGGCTACACCTCCAACTGTAAGGCT 579  
QY 466 cgaatctcgagaagaagtgtctctgccagctgccaccttctctgcccgtagccct 525  
Db 578 CGATCCTCGAGAGAAAGTGTCTGTCGCCAGCTGCCACCTTCTGCGCCGTAGCCCT 519  
QY 526 cactgtgggtgaagacactgagtggtgggaagacacttctgcatgataacagacc 585  
Db 518 CAGTGGGTGTGAAGACCTGGAGTGTGGGAAGGACACTTCTGCCATGATAACAGACC 459  
QY 586 tgcgtccgagacaacacagaggtgtggctgtctgctccatagccacagggctctgtt 645  
Db 458 TGTGTCGAGACAAACGACAGGCTGGGCTGTCTGTCCTACCGCAGCGGCTCTTGT 399  
QY 646 gctgacggtggccactgctgtctgtctgtctgtctgtctgtctgtctgtctgtct 705  
Db 398 GCTGATCGGGCCACTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 339  
QY 706 ttgctgagggagggcccgctgctgggagcccttttggagggagggccctgagagactg 765  
Db 338 TTGCGAGGGAGGCGCCGCGCTGGGAGCGCCCTTTTGGGGAGCCAGCCCTTGAGACAGCTG 279  
QY 766 ctgtgagagacactgaagactgtgagccctgagccctggaggggtgacctc 825  
Db 278 CTGTGAGGGACAGTACTGAAGACTGTGACGCTCTCGGAGCCCTCGGAGGGTGGCCCTC 219  
QY 826 tgcctgagggctctgacactccctcccttaacaaattctccctggagccctctgagct 885  
Db 218 TGTGAGGCTCTCTAGCACTCTCCCTAACCAATCTCCCTGGAGCCCTATCTGAGCT 159  
QY 886 ccccatcacatggaggtgggctccatctaaagccctccctctgagagggggtg 945  
Db 158 CCCCATCACCATGGAGTGGGCTCAATCTAAGG -CTTCCCTGTGAGAGGGGTTG 100  
QY 946 agccaaagcc-cattacaagctgcatccctccctccctccctccctccctccctccag 1004  
Db 99 TGGCAAAAGCCACATTACAAGCTGCAATCCCTCCCGTTTTCAGTGGACCTGTGGCCAG 40  
QY 1005 gtgttttccctacacaggggtgtt 1032  
Db 39 GTGCTTTTCCCTATCCACNAGGTGTT 12

RESULT 3  
AL542659  
LOCUS AL542659 LRI\_FL002\_PL1 Homo sapiens cDNA clone CS0DE014YB02 5 prime  
DEFINITION AL542659, 1 GI:12874922  
ACCESSION AL542659  
VERSION AL542659  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 942)  
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
CONTACT Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
source  
1. 942  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DE014YB02"  
/clone\_lib="LRI\_FL002\_PL1"  
/lab\_host="DH10B"  
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the  
pCMVSPORT 6 vector. Library was constructed by Life  
Technologies. Contact : Feng Liang Life Technologies, a  
division of Invitrogen 9800 Medical Center Drive Rockville  
Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
fliang@lifetech.com URL :  
http://fulllength.invitrogen.com\*  
BASE COUNT 165 a 311 c 287 g 176 t 3 others  
ORIGIN  
Query Match 80.3%; Score 860.8; DB 9; Length 942;  
Best Local Similarity 97.7%; Pred. No. 3.8e-188;  
Matches 901; Conservative 3; Mismatches 15; Indels 3; Gaps 3;  
QY 1 gggcccccacaggtgcccctgagtgagagggcccccagctcacctcagcctgcccagacc 60  
Db 14 GGGCCCCACCCAGGTGCCCTGGATGGAGAGGCCCCAGCTCACCTCAGCTTCCCTGCCAGACCA 73  
QY 61 caagccttgagagagatgtcccctgtgataatgtcagcagctgtccctcctccatcc 120  
Db 74 CAAGCCTTGAAGAGAGATGTCCCTGTGATAATGTACAGCAGCTGTCCCTCCTCCGATACC 133  
QY 121 tgcgtccaaactcagctctgggagtggtggtgctgtccatcccccagaggtgtctgtcgc 180  
Db 134 TGCTGGCAACTCAGCTGTGGGAGTGGGGCTGTCTCCAAATCCCAGAGGCTGTCTGTCTGC 193  
QY 181 tcggaccacagcactgctgccccagcgatcacactgtgtagctgaggggcagtgctcag 240  
Db 194 TCGGACCACCAAGCACTGCTGCCCCAGGGTACACGTGTGTAGCTGAGGGGCAGTGTCTCAG 253  
QY 241 cgaggaaagcagatcgtggtgactggagaagatgctgcccgcggcggttccttatcc 300  
Db 254 CGAGGAAGCGAGATCTGTGGCTGGACTGGAGAAAGATGCTGCCCGCGCGGCTTCTTATCC 313  
QY 301 caccocagagacatcggtgtgacagcacacacagctgcccgtgagggcggaacctgtcgc 360  
Db 314 CACCCAGAGACATCGGGCTGTGACACAGCACACAGCTGCCCGGTGGGCGACACCTTCTGTGC 373  
QY 361 ccgagccaggtggagctggtggtgctgcccagttgcccagtgctgtgtgtgtgtgtgtg 420  
Db 374 CCGAGCCTGGGTGGAGCTGGGCTGTGTCAGTTGCCCATGCTGTGTGTGTGTGTGTGTGT 433  
QY 421 cgcagcactgtgcccggctggctacacctgcaactgcaactgcaactgcaactgcaactg 480  
Db 434 CGCCAGCACTGCTGCCCGGCTGGCTACACCTGCAACGTGAAGGCTCGATCTCGCGAGAG 493  
QY 481 gaagtgtctctccagcctgcccacttctgcccgtagccctcagtggtgtgtgaag 540  
Db 494 GAAGTGGTCTCTGCCAGCCTGCCACCTTCTTGGCCCCGTAGCCCCCTACCTGCTGTGTGT 553  
QY 541 gacgtgagtggtgggaagagacacttctgcatgataaccagacactgctgcccagacac 600  
Db 554 GACCTGAGTGTGGGAGGAGACACTTCTGCCATGATACAGACACCTGCTGCCAGACAA 613  
QY 601 cgacagggctgggctgtgtctccatccagggcggtgtgtgtgtgtgtgtgtgtgtgtgtgt 660  
Db 614 CGACAGGCTGGGCTGTGTCTTCCCTTACCGCCAGGGCGCTGTGTGTGTGTGTGTGTGTGT 673  
QY 661 tgcgtcctgtggtcttcgctgagcagcaggggtacaaagtgtgtgtgtgtgtgtgtgtgtgt 720  
Db 674 TGCTGTCTGTGTCTTCCCTGTCGACAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 732  
QY 721 ccgctggagcagccctttgagggaccagccttgagacagctgtgtgagggagagta 780  
Db 733 CCGCGCTGGGAGCGCCCTTTTGGAGGACCCAGCCCTTGACACAGCTGCTGTGAGGGACAGTA 792  
QY 781 ctgaagactctgcagccctcgtggagcccccactcgtggaggtgtccctctgtcagggccct 840  
Db 793 CTGAAGACTCTGCAGCCCTCGGGAGCCCCACTCGGAGGGTGTCCCTCTGTGTGTGTGTGTGT 852  
QY 841 agcactcccccataacaaattctccctggaccctatctgagctcccccatacattgtgtgtgt 900  
|||||

```

Db 853 AGCACTCCCTTAMCAAAATTCCTCTGGA-CCCAATCTGAGCTCCCATCACCATGGG 911
QY 901 aggtggggcctcaatcaagc 922
Db 912 AAGTGGGG-CTCATCTAGGC 932

RESULT 4
AL574944/c 900 bp mRNA linear EST 16-FEB-2001
LOCUS AL574944 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI064YG14 3
DEFINITION prime, mRNA sequence.
ACCESSION AL574944
VERSION AL574944.1 GI:12935632
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 900)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI064YG14"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 171 a 269 c 290 g 153 t 17 others
ORIGIN

Query Match 76.3%; Score 817.8; DB 9; Length 900;
Best Local Similarity 95.7%; Pred. No. 1.5e-178;
Matches 863; Conservative 16; Mismatches 19; Indels 4; Gaps 4;

QY 80 tccctgtgataatgtcagcagctgtccc-tctccatatactgtcgtcgtcccaactcagctct 138
Db 900 YCCCTGTGATATGTGTCAGCAGCTGTCCGTCCTCCGATACCTGCTGCCAACTCACTGCT 841
QY 139 ggggagtggggctgtctcccaatccccagaggtgtctgtcgtcggaccacagcagctgc 198
Db 840 GGGGAGTGGGGCTGCTGTCCAAATCCAGAGGCTGTCTGTCTGCTCGGACCACCACTGTC 781
QY 199 tggccacagcgtacacagtgtagctgaggggagtgtagcagcagagagagagatcgtg 258
Db 780 TGCCCCCAGGGCTACACGTGTGTAGCTGTAGGGGAGGTGTACGCGAGGAGGAGAGATCGTG 721
QY 259 gctgagctgagagagatgctgcgcgcgcgcgttccttattccaccacccagagacatggc 318
Db 720 GCTGGACTGTGAGAGATAGCTCTG-CCGCGGGCTTCTTATATCCCAACCCAGAGACATGGC 662
QY 319 tgtgaccacacacacagctgcccgggtggcggaacctgctgcccagagccaggggtggagc 378
Db 661 TGTGACCAGCACACACAGCTGCCGGTGGGCGAGACCTGCTGCCCGAGGCTGGGTGGGAGC 602
QY 379 tgggacctgctgccagttgccccatgctgtgtgtgtgagagatgccagcagcagctgctgccg 438

```

```

Db 601 TGGGCTGTCTGTCAGTTGCCCATGTGCTGTGTGTCGAGATGCCAGCACTGCTGCCCG 542
QY 439 gctggctacacctgcaacgtgaaggtcctgctgcgagaggaagtggtctctgcccag 498
Db 541 CTGGGTACACCTGCAACGTGAAGGCTCGATCTCTCGAGAGGAGGTGGTCTCTGCCCG 482
QY 499 cctggcaccttctcctggcccttagccctcactgctggtgtgaagacgtggtggtgaa 558
Db 481 CCTGCCACCTTCTCTGGCCCTAGCCCTACGCTGGGTGTGAAGGACGTGGAGTGTGGG 422
QY 559 ggcacattctgcatataaccagacgtgctgcgagacacacacaggggctgctgc 618
Db 421 GRAMACTTCTGCCATATATACACACCTGTGTCGAGACACACCCAGGGGTGGGCTGC 362
QY 619 tgcctacgcccagggcgtctgtgtgtgctgacgctgctgctgctgctgctgctgct 678
Db 361 TGTCCTTACCGCAGGGCGTCTGTGTGCTGTATCGGGGCCACACCTGCTCTGCTGG 302
QY 679 cgtgctgacgctgaggggtaccaggtgtgtgctgacgctgctgctgctgctgctgct 738
Db 301 CCSTGGCGCAGCCAGGGGTACCAAGTGTTCGCGCAGGAGSASCSCGSGCTGGGACG 242
QY 739 ttgaggacccagccttgagacagctgctgtgagggacagctgctgagacactctgcag 798
Db 241 TTGAGGACCCAGCGCTTGAGACAGCTGCTGTGAGGGACAGTACTGAAGACTCTGCAG 182
QY 799 tcgggacccactcgaggggtgctgctgctgctgctgctgctgctgctgctgctgct 858
Db 181 TCGGGACCCACACTCGGAGGGTGCCTCTGTCTAGGCTCTCCCTAGCACCCTCTCCST 122
QY 859 aattctcctggaccccttctgagctcctcctcctcctcctcctcctcctcctcctcct 918
Db 121 AATTCTCCTTGGACCCCATTTCTGAGCTCCCTCCCTCCTCCTCCTCCTCCTCCTCCT 62
QY 919 aggccttctcctgctcagaaggggttgagcagaagcc-cattacaagctcctcctcct 977
Db 61 AGG-CCCTTCTCCTGTGAGAGGGGGTGTGGCAAAAGCCACATTAAAGCTCCCATCC 3
QY 978 cc 979
Db 2 CS 1

RESULT 5
AL576991/c 887 bp mRNA linear EST 16-FEB-2001
LOCUS AL576991 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI081YC03 3
DEFINITION prime, mRNA sequence.
ACCESSION AL576991
VERSION AL576991.1 GI:12939682
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI081YC03"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end

```



```
Db 242 GAGATCGTGGCTGGACTGGAGAGATGCTGCCCGCCGGGCTTCCTTATCCACCCAG 301
Qy 310 gacatcgctgtgaccagcacaccagctgcccgggtggcggaacctgtgcccagcag 369
Db 302 GACATCGGCTGTGACCAAGACACAGCTGCCGGTGGGGCAGACCTGCTGCCCGAGCCGTG 361
Qy 370 ggtggagctgggctgtgcccagttgcccacatgctgtgtgtgctgagagatcgccagcac 429
Db 362 GGTGGAGCTGGGCTGCTGCCAGTGGCCATGCTGTGTGCTGCGAGGATGCCAGCAC 421
Qy 430 tctctccgctgtgtatatacctgaacctgaacctgaagctgacatcctgagagaagtgctc 489
Db 422 TGTGCCCCGGCTGGGTACACCTGCAACGCTGAAGGCTCGATCTGCGAGAAGAGTGTCTC 481
Qy 490 tctgcccagctgcccacttcttctgcccagtagccctcacgtgggtgtgaagagctgag 549
Db 482 TCTGCCAGCCTGCCACCTTCTGCCCCGTACCCCTCAGCTGGGTGTGAGAGACGTGGAG 541
Qy 550 tgtgggaagagacattcttgcattgataaccagacctgtgctgcgagacaaacagcagggc 609
Db 542 TGTGGGAGGAGACACTTCTGCCATGATACACAGACTGCTGCCGAGACAAACGACAGGCG 601
Qy 610 tgggctgctgctccctacagccagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 669
Db 602 TGGGCTGCTGCTTACCCGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
Qy 670 gctggtctccgtgcgacgaggggtaccagtgctgttcgagggagggcccgctgag 729
Db 662 GCTGGCTTCCGCTGGCAGCCAGGGGTACCAAGTGTATGCGAGGAGGCCCGG-GCTGG 720
Qy 730 gacgccccttgaggagaccagccttgagacagctgtgtgtgtgtgtgtgtgtgtgtgtgt 789
Db 721 GACGCCCTTTGAGGGACCCAGGCTTGACAGAGCTGCTGTGAGGGACACTACTGAAGACT 780
Qy 790 ctgacgacctgggacct 807
Db 781 CTGACGCCCTCGGGAGCCC 798

RESULT 7
AL562633/c
LOCUS AL562633 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC015YA01 3
DEFINITION prime, mRNA sequence.
ACCESSION AL562633
VERSION AL562633.1 GI:12911246
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 908)
Li W.B., Gruber C., Jessee J., and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC015YA01"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
```

```
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com*
BASE COUNT 153 a 292 c 296 g 157 t 10 others
ORIGIN
Query Match 70.6%; Score 757; DB 9; Length 908;
Best Local Similarity 93.3%; Pred. No. 1.6e-164;
Matches 847; Conservative 6; Mismatches 49; Indels 6; Gaps 6;
Qy 98 gagctgttctctccctccgataacctgtg-ccaactcaagctggggagtggggctgctgt 156
Db 908 GGGGCTGTCCCTCCCTCCGATACCCGCTGCCCAATACACTCTGGGAGTGGGGCGGCTGT 849
Qy 157 ccaa-tcccagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 215
Db 848 CCAATCCAGAGGCTGTGTGCTGCTGCGACACAGACAGGCTGCCCGGCAACAC 789
Qy 216 gtgttagctgagggcagctgtcagcaggaagcagatcgttgctggaactgagaagat 275
Db 788 GTGTGTAGCTGAGGGCAGTGTGACGGAGGAAGCGAGATCGTGGCTGAAGTGGAGAGAT 729
Qy 276 gctgtccgctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 335
Db 728 GCYTSGCCCGGGCTTCCTTATCCACCAAGARACATCGGCTGTGACGACACACACAG 669
Qy 336 ctgcccgtgtggcgaacctgtgcccagcaggggtgggagctgggctgtgctgacctt 395
Db 668 CTGCCCGTGGGGCAGACCTGCTGCCGAGCTGGGTGGGAGCTGGGCTGTGTGCCAGTT 609
Qy 396 gcccagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 454
Db 608 GCCCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 549
Qy 455 acgtgaagctcgatctctgcgagaaggaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 514
Db 548 ACCTGAAGGCTCGATCTCGAGGAAGAAAGTGTCTCTGCCAGGCTGCCACCTNCTGG 489
Qy 515 cccgtagcctcacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 574
Db 488 CCCGTAGCCCTCACCTGGGTGTGAAGAGCTGTGGGAGAGACACTTCTTGCCTATG 429
Qy 575 ataacagacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 634
Db 428 ATAACAGACCGCTGTGCGAGACACCGAGAGGCTGGGCGGCTGTCTCTACCGCCAGG 369
Qy 635 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 694
Db 368 GCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 309
Qy 695 gtacaaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 754
Db 308 GTACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 249
Qy 755 tgagacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 814
Db 248 TGAGAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189
Qy 815 aggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 874
Db 188 AGGGTGGCTTCNGCTCAGGCGCTCCCTAGCACCTCTCCCTTAACCAATTTCTCCTGGACCC 129
Qy 875 cattctgagctcccatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 934
Db 128 CATTCTGAGCTCCCATCACCATGGGAGTGGGGGCTCAATCTAAGG-CTTTCCTGTGCA 70
Qy 935 gagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 992
Db 69 GAAGGGGGGATGTGGCAAGACCAACAGAGCTGCCATGCCCTCCCTCGCCGCTGTCTAGTGTGR 10
```







[illegible]

```

FEATURES
source
1. 807
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3836510"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="organ: ovary; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GCACACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
135 a 259 c 258 g 155 t
BASE COUNT
ORIGIN

```

Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAMI1504 row: h column: 19  
 High quality sequence start: 23  
 High quality sequence stop: 830.

```

FEATURES
source
1. 894
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5201346"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site1: NotI; Site2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-d-
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
168 a 274 c 293 g 159 t
BASE COUNT
ORIGIN

```

[illegible]

Qy	138	tgaggagtggggtgctgtccaatcccagaggtgtctgct-gctcggaccaccagcact	196
Db	894	TCGGGACTGGGGTGTGCTGCCAATCCAGAGCTGTCGGCTGCTCGGACCACGACCT	835
Qy	197	gctgccccccagcagatac-acgtgtgtagctgaggggcagtgctcagcgagaga-agcgagat	254
Db	834	GCTGCCCCAGGGGCACAACAGCTGTGAAGCTGAGGGGCACCTGTTCAGCGAGGAGAGCGAGAT	775

QY 255 cgtgctggac--tggagaagatgctcccgccgctgcttcttattccacccacccagagac 312  
|||||  
Db 774 CTTGCTGGACCTGGAGATAATGCTGCGCCGCGGCTTCTGATCCACCCAGAGAC 715  
QY 313 atcgctgtgaccagacacacagctgcccgtggcgga-acctctgccagccaggg 371  
|||||  
Db 714 ATCGCTGTGACACACACAGCTGCGCGGTAGGGGAGAGCTGCTGCCGAGCGTGG 655  
QY 372 tggagctgggctg-ctgceagttgccccatgctgtgtgctgagagatgccagac- 429  
|||||  
Db 654 TGGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595  
QY 430 tctgccccgc-tggctacacctgcaacgtgaagctgctgctgctgctgctgctgctgct 488  
|||||  
Db 594 TCTGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535  
QY 489 ctctccagctgcaaccttctgctgctgctgctgctgctgctgctgctgctgctgctgct 548  
|||||  
Db 534 CTCTGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 475  
QY 549 gtgtggggaagacacttctgcatgataacacagacactgctgctgctgctgctgctgctgct 608  
|||||  
Db 474 GTGTGGGGAAGACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 415  
QY 609 ctgggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 668  
|||||  
Db 414 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 355  
QY 669 tctggtctccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 728  
|||||  
Db 354 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295  
QY 729 gtagccctctttagggagccacagctttagagagctgctgctgctgctgctgctgctgctgctgct 788  
|||||  
Db 294 GAGGCGCTTGTAGGAGCCAGCTTGTAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235  
QY 789 tctgagccctgggagccacactgaggggtgctgctgctgctgctgctgctgctgctgctgctgct 848  
|||||  
Db 234 TCTGACGCTCGGGACCCACTCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175  
QY 849 cccctaacaattctccctgagccacattctgagctcc-ccataccatggaggtggg 907  
|||||  
Db 174 CCCTTAACCAATTCCTCTGGACCCCACTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 115  
QY 908 gctcaatcaagcccttccctgctgagaggggtgaggggaggggaggggaggggaggggagggg 966  
|||||  
Db 114 GCCTCAATCTAAGG-CCTTCCCTGTGAGAGGGGGTGTGGCAAGGCCACATTACAAGC 56  
QY 967 tgcataccctccctcccttccagtgacc 994  
|||||  
Db 55 TGCCATCCCTCCCTCCCTTTCAGTGGACC 28

RESULT 14  
BG830299  
LOCUS 602764696F1 NIH\_MGC\_42 Homo sapiens CDNA clone IMAGE:4906991 5',  
DEFINITION mRNA sequence.  
ACCESSION BG830299  
VERSION BG830299.1 GI:14177886  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 751)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM1810 row: o column: 24  
High quality sequence stop: 748.  
FEATURES  
Location/Qualifiers  
source  
1..751  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4906991"  
/clone\_lib="NIH\_MGC\_42"  
/tissue\_type="epitheloid carcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;  
site: 2: EcoRI; cDNA made by oligo-dT priming.  
directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library. |"  
BASE COUNT 125 a 249 c 227 g 150 t  
ORIGIN

Query Match 62.8%; Score 673.2; DB 10; Length 751;  
Best Local Similarity 96.5%; Pred. No. 3.4e-145;  
Matches 720; Conservative 0; Mismatches 23; Indels 3; Gaps 3;  
QY 286 cgcggtcttctatccacccacagagacatcggtgtgaccagcacacacagctgcccggtg 345  
Db 2 CGGGCTTCTTATCCACCCACAGAGACATCGCTGTGACGACACACAGCTGCCGGTG 61  
QY 346 ggcggaaactgctgcccagagccaggggtggagctgggctgctgctgctgctgctgctgctgct 405  
Db 62 GGGCAGACCTGCTGCCGAGCCTGGGTGGAGGTGGGCCCTGCTGCCAGTGTGCCCATGTCT 121  
QY 406 gtctgctgagagatgcacagcagctgctgcccggctgctacacctgcaacgtgaaagct 465  
Db 122 GTGTGCTGCGAGAGATCGCCAGCACTGCTGCCGGCTGGCTACACCTGCAACGTGAAGCT 181  
QY 466 cgaacctgcagaggaagtggtctctgcccagcctgcccacctctctctggtgcccctagccct 525  
Db 182 CGATCCTGCGAGAGGAAGTGGTCTCTGCGCAGCCTGCCACCTTCTCTGGCCGCTAGCCCT 241  
QY 526 cactgggtgtgaaggacgtggaggtggggaaggagacacttctgccaatgataccagacc 585  
Db 242 CACGTGGGTGTGAAGGACGTGGAGTGTGGGAAGGACACTTCTGCCATGATAACACAGACC 301  
QY 586 tgcctgcagacacacagcagggctgctgctgctgctgctgctgctgctgctgctgctgctgct 645  
Db 302 TGTGCGAGACACACACAGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361  
QY 646 gct 705  
Db 362 GCTGATCGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 706 ttgcaggggagggcccgctggagcccttggagggagccagccttggagagcagctg 765  
Db 421 TTGGCAGGGAGGCGCGCTGGGCGCCCTTTTGGAGGACCCAGCCTTTGAGACAGCTG 480  
QY 766 ctgtgagggagcagctgagagctgctgctgctgctgctgctgctgctgctgctgctgctgct 825  
Db 481 CTGTGAGGGAGACAGTACGAGACTCTGCAGCCCTCGGAGACCCCACTCGGAGGGTGCCTC 540  
QY 826 tgcagggctccctagcactccctcccttaacaaattctcctgagcccttctgagct 885  
Db 541 TGCTCAGGCTCCCTAGCACCCTCCCTTACCAATTTCTCCTGGACCCCACTTCTGAGCT 600  
QY 886 ccccatcacatgggaggtggggtcctcaatcaatgaaggccttccctgctgcaagaagggttg 945





Result No.	Score	Query Match	Length	DB	ID	Description
1	1072	100.0	2095	4	US-08-991-862-16	Sequence 16, Appl
2	747.2	69.7	1779	1	US-07-668-648-3	Sequence 3, Appl
3	747.2	69.7	1779	2	US-08-429-998-3	Sequence 3, Appl
4	747.2	69.7	1779	2	US-08-431-333-3	Sequence 3, Appl
5	747.2	69.7	1779	5	PCR-US91-02321-3	Sequence 3, Appl
6	533.8	49.8	2137	4	US-08-991-862-1	Sequence 1, Appl
7	469.2	43.8	1767	1	US-07-668-648-1	Sequence 1, Appl
8	469.2	43.8	1767	2	US-08-429-998-1	Sequence 1, Appl
9	469.2	43.8	1767	2	US-08-431-333-1	Sequence 1, Appl
10	469.2	43.8	1767	5	PCR-US91-02321-1	Sequence 1, Appl
11	448.4	41.8	1767	1	US-07-668-648-5	Sequence 5, Appl
12	448.4	41.8	1767	2	US-08-429-998-5	Sequence 5, Appl
13	448.4	41.8	1767	2	US-08-431-333-5	Sequence 5, Appl
14	448.4	41.8	1767	5	PCR-US91-02321-5	Sequence 5, Appl
15	153.2	14.3	539	1	US-07-668-648-7	Sequence 7, Appl
16	153.2	14.3	539	2	US-08-429-998-7	Sequence 7, Appl
17	153.2	14.3	539	2	US-08-431-333-7	Sequence 7, Appl
18	153.2	14.3	539	5	PCR-US91-02321-7	Sequence 7, Appl
19	64.6	6.0	341	1	US-07-668-648-9	Sequence 9, Appl
20	64.6	6.0	341	2	US-08-429-998-9	Sequence 9, Appl
21	64.6	6.0	341	2	US-08-431-333-9	Sequence 9, Appl
22	64.6	6.0	341	5	PCR-US91-02321-9	Sequence 9, Appl
23	37	3.5	20303	1	US-08-370-975B-6	Sequence 6, Appl
24	37	3.5	26764	1	US-08-370-975B-1	Sequence 1, Appl
25	36.2	3.4	4403765	4	US-09-103-840A-2	Sequence 2, Appl
26	36	3.4	935	2	US-08-822-038-73	Sequence 73, Appl
27	36	3.4	935	4	US-08-479-285-73	Sequence 73, Appl



Qy	601	cgacagggtggtgctgtctctatcgcccccagggtctctgttctgtgtagtcgagcgccac	660
Db	1612	CGACAGGGCTGGGCGCTGTGTCCTTACCGCCAGGCGGTCTTTGTGTGTATCGGCGCCAC	1671
Qy	661	tgcgtctctgtggtcttcgctgcgcacgcgaggggtaccaaagtgttttcgcagaggagggcc	720
Db	1672	TGCTGTCTCTTGGCTTTCGCTTCGCAGCCAGGGGTACCAAGTGTTTTCGCAGGAGGCC	1731
Qy	721	cgcgcgtggagcgcgccctttgaggagaccagcccttgacacagctgcgtg	768
Db	1732	CCGCGCTGGGAGCGCCCTTTGAGGAGCCAGCGCTTGACACAGCTGCTG	1779

RESULT 3  
 US-08-429-998-3  
 ; Sequence 3, Application US/08429998  
 ; Patent No. 5885961  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shoyab, Mohammed  
 ; APPLICANT: Plowman, Gregory D.  
 ; TITLE OF INVENTION: EPITHELINS: NOVEL CYSTEINE-RICH GROWTH  
 ; TITLE OF INVENTION: MODULATING PROTEINS  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/429,998  
 ; FILING DATE: 27-APR-1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/668,648  
 ; FILING DATE: 13-MAR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mistock, S. Leslie  
 ; REGISTRATION NUMBER: 18,872  
 ; REFERENCE/POCKET NUMBER: 5624-161-999  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)790-9090  
 ; TELEFAX: (212) 869-9741  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1779 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 ; TISSUE TYPE: Kidney  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..1779  
 ; US-08-429-998-3

	Query Match	69.7%	Score 747.2;	DB 2;	Length 1779;
	Best Local Similarity	98.3%;	Pred. No. 2.9e-175;		
	Matches 755;	Conservative	0;	Mismatches 13;	Indels 0;
	Gaps	0;			
Qy	1	ggggcccccaccagggtgcctctgattgagaagagcccccagctcacctcagctgcagaccga	60		
pb	1012	ggggcccccaccagggtgcctctgattgagaagagcccccagctcacctcagctgcagaccga	1071		

Qy	61	caagccttgaagagagatgtcccctgtgtataatgtcaagcagctgtccctctccatatacc	120
Db	1072	CAAGCCTTTGAAGAGAGATGTCCCTGTGTATATGTACAGAGCTGTCCCTTCCCGCATAC	1131
Qy	121	tgtctgccaactcagctctggggagtgggctgtctgtccaatcccaagagctgtctgctgc	180
Db	1132	TGCTGCCAACTCACGTCTGGGGAGTGGGGCTGCTGTCCAAATCCACAGAGGCTGTCTGTGTCG	1191
Qy	181	tcggaccaccagcactgtgtccccccagcgatatacacgtgtgtagtctgagggggcagtgctcag	240
Db	1192	TCGGACCACCAAGCACTGCTGTGCCCCCAGGGCTACACGTGTGTACCTGAGGGCGAGTGTCCAG	1251
Qy	241	cggagaaacgagatcgtgtgcttgagctgagaaagatgcttgcctgccgcgcggttctcttatcc	300
Db	1252	CGAGGAACGCGAGATCGTGGCTGGACTGGAGAAGATGCTCTGCCGCCGCCGGCTTCTTTATCC	1311
Qy	301	caccccagagacatcggctgtgaccagcacacacagctgccccggtggggcgaaacctgtctgc	360
Db	1312	CACCCAGAGACATCGGCTGTGACCAGCACACACAGTGCCTGCCGCTGGGGCAGACCTGCTGC	1371
Qy	361	ccgagccagggtggagatgggctgtctgccaattgtccccaatgctgtgtgtctgcgaagat	420
Db	1372	CCGAGCCCTGGGTGGAGCTGGGCCCTGCTGCCAGTTGCCCCATGCTGTGTGCTGCGAGGAT	1431
Qy	421	cgcacgactgctgcccggtggctacacctgcaacgtgaaaggtctgatacctctgcgagaag	480
Db	1432	CGCCAGCACTGCTGCCCGGCTGGCTACACCTGCAAGCTGAAGGCTCGATCCTGCGAGAAG	1491
Qy	481	gaagtgtctctgccagcctgccacctctctggccgttagccctcaatgggtgtggaag	540
Db	1492	GAAGTGGTCTCTTGCCACAGCTGCCACCTTCTTCCCTGGCCCGTAGCCCTCACGCTGGGTGTGAAG	1551
Qy	541	gacgtggagtgggaagacacitctccatataaccagacctgtctgcgcagacaac	600
Db	1552	GAGCTGGAGTGTGGGAAGGACATCTCTCCATGATAACCAAGACTGCTGTCCGAGACCAAC	1611
Qy	601	cgcacgggtggcgctgtctccctaccgcccgaggcgctgtgtgtgtctgatoaggccac	660
Db	1612	CGACAGGGCTGGCCCTGCTGTCCCTACCGCCAGGGGGCTGTGTTGTGCTGATCGCGCGCAC	1671
Qy	661	tgtctctctgtgtggtcttcgctgcgcacgcaggggtaccaagtgtttgcgcaggagccc	720
Db	1672	TGCTGTCTCTGTCTGGCTTTCGCTGCGCAGCGAGGGGTACCAAGTGTGTTTGCACGGGAGGCC	1731
Qy	721	ccgcgtggagcgcctcttgagggaccagccttgagacagctgtgtg	768
Db	1732	CTGCGCTGCGGACGCCCCCTTTTGAGGACCCAGCCTTGAGACAGCTGTGTG	1779

RESULT 4  
US-08-431-333-3  
Sequence 3, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plovman, Gregory D.  
TITLE OF INVENTION: EPIPELINS: NOVEL CYSTEINE-RICH GROWTH  
PEPTIDES  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Invention #1.0, Version #1.25  
CURRENT APPLICATION DATA:







LENGTH: 1767 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1767  
US-08-429-998-1

Query Match 43.8%; Score 469.2; DB 2; Length 1767;  
Best Local Similarity 76.7%; Pred. No. 8.6e-107;  
Matches 589; Conservative 0; Mismatches 173; Indels 6; Gaps 1;  
QY 1 gggccccaccaggtgcccctgagtgagagagggccccagctcacctcagcctgcccagaccga 60  
DB 1006 GGAGTCTCTCAGGTACCTGGATGAAAAGGTCAAGGCTCCCTCAGCCTGCCAGACCCA 1065  
QY 61 caagccttgaagagagatgccctgtgataatgtcagcagctgcccctcccgatacc 120  
DB 1066 CAGATCTTGAAGATGATGCTCCCTGTGATGACTTCAAGTGTCTTCAACAATACC 1125  
QY 121 tgcctgcaactcacgtctggggagtgggctgtgtcccaatccccagaggtgtgtctgtgc 180  
DB 1126 TGCTGCACACTCAGTCTTGGGGACTGGGGCTGTCTCCATCCAGAGGCTGTCTGTGTC 1185  
QY 181 tcggaccaccagcactgctgcccagcagcgatatacacgtgtgttagctgagggcgagtgc 240  
DB 1186 TTAGACCACCAGCATGTGCTGCCCTCAGGCTTTCAAAATGTATGATGAGGGGTACTGTAG 1245  
QY 241 cgaggaagcagagatcgctggctggagtgagagagatgcctgcccgcgcgcgttccctatcc 300  
DB 1246 AAGGGAGACAGAAATGCTGGCTGGCTGGAGAGATGCTGTCCGCCAGACAACTCTGTCTC 1305  
QY 301 caccaccagagacatcggtctgaccagcacaccagctgcccgggtggcgagacgtgtgc 360  
DB 1306 CAACATGGAGATATTGTTGTGACAGCATACACAGCTGCCAGTAGGGCAACATGCTGTC 1365  
QY 361 ccgagcaggggtggagctggcctgctgcccagctgcccagctgctgtgtgtgcgagat 420  
DB 1366 CCAAGCCTGAAGGAATGCTGGGCTGTGCTGAGGATGCTGTGCTGTGCTGTGAGGAC 1425  
QY 421 cgccagcactgctcccgctggctacacctgcaacgtgaaagctcgatccctgcgagag 480  
DB 1426 CGGAGGACTGTTCCCGGCTGGGTACACTGCAACGTGAAGGCGAGAACCTGTGAGAG 1485  
QY 481 gaagtgtctctgcccagcctgcccacttccctgcccagtagccctcagctgggtgtgag 540  
DB 1486 GATCAGGCTCTGTCCAGCCTTCCATGGACCTGACCTTTGGCTCTAAGGTTGG-----G 1539  
QY 541 gacgtgagtggtgggaaggaacattctgcatataacagacactgctgcccagacac 600  
DB 1540 AATGTGAATGTGTGGCGGACATTTCTGCCATGATAACCACTGCTGTGTAAAGACAGC 1599  
QY 601 cgacagggctggcctgctgcccagcagggcgtctgtgtgtgtgtgtgtgtgtgtgtgt 660  
DB 1600 CAAGGAGCTGGGCTGCTGCTCCCTATGTAAGGGGTGCTGCTGTGAGAGATGGAGTCCAC 1659  
QY 661 tgcctgctgctgcttccgtgctgcagcaggggtgtacaaagtgtttgtcagggagggc 720  
DB 1660 TGTGTCCTCATGCTTCCACTGTTTACGCCAAGGGAACCAAGTGTGCGGAAGAAGACC 1719  
QY 721 ccgctgagggacgcccctttgagggaccagccttgagacagcgtgtg 768  
DB 1720 CCTGCTGGGACATACTTTTGGGGATCCAGCCCCCAGACCGCTACTG 1767

## RESULT 9

US-08-431-333-1

; Sequence 1, Application US/08431333

; Patent No. 5965723

GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
APPLICANT: Plowman, Gregory D.  
TITLE OF INVENTION: EPITHELIALS: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1767 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1767  
US-08-431-333-1

Query Match 43.8%; Score 469.2; DB 2; Length 1767;  
Best Local Similarity 76.7%; Pred. No. 8.6e-107;  
Matches 589; Conservative 0; Mismatches 173; Indels 6; Gaps 1;  
QY 1 gggccccaccaggtgcccctgagtgagagagggccccagctcacctcagcctgcccagaccga 60  
DB 1006 GGAGTCTCTCAGGTACCTGGATGAAAAGGTCAAGGCTCCCTCAGCCTGCCAGACCCA 1065  
QY 61 caagccttgaagagagatgccctgtgataatgtcagcagctgcccctcccgatacc 120  
DB 1066 CAGATCTTGAAGATGATGCTCCCTGTGATGACTTCAAGTGTCTTCAACAATACC 1125  
QY 121 tgcctgcaactcacgtctggggagtgggctgtgtcccaatccccagaggtgtgtctgtgc 180  
DB 1126 TGCTGCACACTCAGTCTTGGGGACTGGGGCTGTCTCCATCCAGAGGCTGTCTGTGTC 1185  
QY 181 tcggaccaccagcactgctgcccagcagcgatatacacgtgtgttagctgagggcgagtgc 240  
DB 1186 TTAGACCACCAGCATGTGCTGCCCTCAGGCTTTCAAAATGTATGATGAGGGGTACTGTAG 1245  
QY 241 cgaggaagcagagatcgctggctggagtgagagagatgcctgcccgcgcgcgttccctatcc 300  
DB 1246 AAGGGAGACAGAAATGCTGGCTGGCTGGAGAGATGCTGTCCGCCAGACAACTCTGTCTC 1305  
QY 301 caccaccagagacatcggtgtgacacagcacaccagctcccgggtggcgagacgtgtgc 360  
DB 1306 CAACATGGAGATATTGTTGTGACAGCATACACAGCTGCCAGTAGGGCAACATGCTGTC 1365





Query Match 41.8%; Score 448.4; DB 2; Length 1767;  
Best Local Similarity 75.0%; Pred. No. 1.1e-101;  
Matches 576; Conservative 0; Mismatches 186; Indels 6; Gaps 1;

QY 1 gggcccccacagctgctggatggagagggcccccagctcacctcagctgcccagaccac 60  
DB 1006 GGTATCTCCAAAGTACCTGGATGAAGAGGTCATAGTCCCGCCGCGCTGCCAGACCA 1065

QY 61 caagccttgaagagagatgccccctgtgataatgtcagcagctgtccctcctcgtacc 120  
DB 1066 CAGATCTGAAGAGTATACACTTGTGTGATGACTTCACTAGGTGTCTTACAACAATACC 1125

QY 121 tgcgtccaaactcagctgctggagtggtgctgtctgtccaaatccagagctgtctctgc 180  
DB 1126 TGTGCGAAACTCAATCTGGGAGTGGGCTGTCTGCCATCCAGAGGCTGTCTGTGC 1185

QY 181 tcggaccaccagcactgtgctgcccagcagcagcagcagctgtgtagctgagggcagtgacg 240  
DB 1186 TCAGACAACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1245

QY 241 caagccttgaagagagatgccccctgtgataatgtcagcagctgtccctcctcgtacc 300  
DB 1246 AAGGGAGACACAAATGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 1305

QY 301 caccaccagagacatcggtgtgaccagcagcagcagcagcagcagcagcagcagcagc 360  
DB 1306 CAAATTTGGAGATATCGGTTGTGTGACAGCATACCAAGTGGCTGGCTGGCTGGCTG 1365

QY 361 ccgagccaggggtgggagctgctgctgctgctgctgctgctgctgctgctgctgctg 420  
DB 1366 CCAAGCCTCAAGGGAAGTTGGGCTGCTGCCAGCTGCCCATGCTGTGTGTGTGTGTG 1425

QY 421 cgccagcactgtgcccggctggtggtggtggtggtggtggtggtggtggtggtggtg 480  
DB 1426 CGCAGCAGCTGTGTGGCGCGCGGGTACACTGCAATGTGAAGCGGAGGACCTGTGAGA 1485

QY 481 gaagtggtctctgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540  
DB 1486 GATGTCGATTTATTCAGCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1539

QY 541 gacgtggagtggtgggagagacacttctgcccagcagcagcagcagcagcagcagcagc 600  
DB 1540 AATGTGGAGTGTGGAGAGGCAATTTCTGCCATGATAACCAAGACCTGTGTAAGACAGT 1599

QY 601 cgacagggctggcctgctgtccctcagcagcagcagcagcagcagcagcagcagcagc 660  
DB 1600 GCAGGAGTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1659

QY 661 tgcgtcctgctggttccgctgctgctgctgctgctgctgctgctgctgctgctgctg 720  
DB 1660 TGTGCGCCCGTGGCTTCCACTGTTTCAGCCAGGGGAACCAAGTGTGTAAGAGAGATT 1719

QY 721 ccgctgctggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 768  
DB 1720 CCTCGCTGGGACATGTTTTTGGAGGATCCGGTCCCAAGACCGCTACTG 1767

RESULT 13  
US-08-431-333-5  
Sequence 5, Application US/08431333  
Patent No. 5965723  
GENERAL INFORMATION:  
APPLICANT: Shoyab, Mohammed  
TITLE OF INVENTION: EPIITHELIN: NOVEL CYSTEINE-RICH GROWTH  
TITLE OF INVENTION: MODULATING PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,333  
FILING DATE: 27-APR-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/668,648  
FILING DATE: 13-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1767 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1767  
US-08-431-333-5

Query Match 41.8%; Score 448.4; DB 2; Length 1767;  
Best Local Similarity 75.0%; Pred. No. 1.1e-101;  
Matches 576; Conservative 0; Mismatches 186; Indels 6; Gaps 1;

QY 1 gggcccccacagctgctggatggagagggcccccagctcacctcagctgcccagaccac 60  
DB 1006 GGTATCTCCAAAGTACCTGGATGAAGAGGTCATAGTCCCGCCGCGCTGCCAGACCA 1065

QY 61 caagccttgaagagagatgccccctgtgataatgtcagcagctgtccctcctcgtacc 120  
DB 1066 CAGATCTGAAGAGTATACACTTGTGTGATGACTTCACTAGGTGTCTTACAACAATACC 1125

QY 121 tgcgtccaaactcagctgctggagtggtgctgtcgtccaaatccagaggtgtctgtctgc 180  
DB 1126 TGTGCGAAACTCAATCTGGGAGTGGGCTGTGTGCCATCCAGAGGCTGTCTGTGC 1185

QY 181 tcggaccaccagcactgtgctgcccagcagcagcagcagcagcagcagcagcagcagc 240  
DB 1186 TCAGACAACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1245

QY 241 caagccttgaagagagatgccccctgtgataatgtcagcagctgtccctcctcgtacc 300  
DB 1246 AAGGGAGACACAAATGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 1305

QY 301 caccaccagagacatcggtgtgaccagcagcagcagcagcagcagcagcagcagcagc 360  
DB 1306 CAAATTTGGAGATATCGGTTGTGTGACAGCATACCAAGTGGCTGGCTGGCTGGCTG 1365

QY 361 ccgagccaggggtgggagctgctgctgctgctgctgctgctgctgctgctgctgctg 420  
DB 1366 CCAAGCCTCAAGGGAAGTTGGGCTGTCTGCCAGCTGCCCATGCTGTGTGTGTGTGTG 1425

QY 421 cgccagcactgtgcccggctggtggtggtggtggtggtggtggtggtggtggtggtg 480  
DB 1426 CGCAGCAGCTGTGTGGCGCGCGGGTACACTGCAATGTGAAGCGGAGGACCTGTGAGA 1485



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 5, 2002, 16:30:39 ; Search time 3362.64 Seconds  
(without alignments)  
6671.320 Million cell updates/sec

Title: US-09-824-647-16\_COPY\_1024\_2095  
Perfect score: 1072  
Sequence: 1 gggccaccaggcgccg.....ataaagttgtcaacttttt 1072

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	GenEmbl:*
1:	gb_ba:*
2:	gb_hlg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vi:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_vi:*
30:	em_htg_hum:*
31:	em_htg_inv:*
32:	em_htg_other:*
33:	em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description

1	1072	100.0	2095	6	ARI75456
2	1072	100.0	2095	9	HUMGRANUL
3	992.6	92.6	2132	9	BC010577
4	992.6	92.6	2152	9	HSEPTT1
5	992.6	92.6	2178	9	AF055008
6	991	92.4	2157	9	AK000607
7	989.4	92.3	1630	9	AK023348
8	747.2	69.7	1779	6	AR079655
9	747.2	69.7	1779	6	II1830
10	629.2	58.7	139488	9	AC003043
11	621.4	55.6	180520	9	HUMGRANULI
12	596.4	55.6	180520	2	AC019152
13	565	52.7	189391	2	AC019151
14	560.6	52.3	2153	10	RNEPIT1
15	557.4	52.0	2113	10	RATGABE12A
16	546.4	51.0	2125	10	MUSAG
17	533.8	49.8	2137	6	ARI75447
18	532.2	49.6	2100	10	MMEPIT1
19	492	45.9	1867	9	BC000324
20	469.2	43.8	1767	6	AR079654
21	469.2	43.8	1767	6	II1829
22	448.4	41.8	1767	6	AR079656
23	448.4	41.8	1767	6	II1831
24	364	34.0	55774	2	AC091303
25	294.4	27.5	1952	10	GPIAG
26	267.8	25.0	8094	10	MUSAP
27	255.4	23.8	165490	2	AC025326
28	255.4	23.8	220710	2	AL596258
29	197.8	18.5	336	6	AX198393
30	197.8	18.5	336	6	AX208943
31	153.2	14.3	539	6	AR079657
32	153.2	14.3	539	6	II1832
33	111.4	10.4	113	11	G54019
34	64.8	6.0	125020	9	AF429315
35	64.6	6.0	341	6	AR079658
36	64.6	6.0	341	6	II1833
37	63.6	5.9	125020	9	AF429315
38	61.2	5.7	18347	3	CET2H2
39	56.6	5.3	135369	2	AC090557
40	56.6	5.3	173310	2	AC090454
41	56.6	5.3	181923	9	AC022872
42	44	4.1	110000	2	LMFLCHR18_08
43	43.4	4.0	131370	30	AC017018
44	43.4	4.0	194236	2	AL645944
45	43.2	4.0	202000	1	AP000058

ALIGNMENTS

RESULT 1	ARI75456	2095 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	ARI75456	Sequence 16 from patent US 6309826.			
DEFINITION	ARI75456	Sequence 16 from patent US 6309826.			
ACCESSION	ARI75456				
VERSION	ARI75456.1	GI:17916755			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2095)				
AUTHORS	Serrero, G.				
TITLE	88kDa tumorigenic growth factor and antagonists				
JOURNAL	Patent: US 6309826-A 16 30-OCT-2001;				
FEATURES	Location/Qualifiers				
source	1. 2095				
BASE COUNT	364 a 682 c 624 g 425 t				
ORIGIN	/organism="unknown"				

Query Match 100.0%; Score 1072; DB 6; Length 2095;  
Best Local Similarity 100.0%; Pred. No. 2.6e-234;

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/668,648  
FILING DATE: 19910819  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-161-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212) 869-9741  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Bos taurus  
TISSUE TYPE: Kidney  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..537  
US-07-668-648-7

Query Match 14.3%; Score 153.2; DB 1; Length 539;  
Best Local Similarity 84.0%; Pred. No. 4.3e-29;  
Matches 173; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 1 gggccccaccaggtgccctggatggagagggccccagctcacctcagcctgcagacccca 60  
DB 328 gggacccgccaggtgccctggatgaaagaaagcccccagccctcagcctgcagacccca 387  
QY 61 caagccttgaagagagatgccctgtgataatgtcagcagctgtccctccctccgatacc 120  
DB 388 ggagcagtgagggggagcgtccctgtgataagctcaccagctgtcttccactacc 447  
QY 121 tgcgtccaaactcacgtctgtggagtggtggtgtgtcccaatccagaggtgtctgtgc 180  
DB 448 tgcgtgcgactcaagctcggggagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgc 507  
QY 181 tcggaccaccagcactgtgccccca 206  
DB 508 tcggaccaccagcactgtgtcccca 533

Search completed: September 5, 2002, 18:22:29  
Job time: 13352 sec

```
Matches 1072; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gggccccaccagtgccctggatggagagggcccccagctcacctcagctgagccagaccca 60
DB 1024 GGGCCCCCACCAGTGCCCTGGATGGAGAGGGCCCCAGCTCACCTCAGCTGCCAGACCCA 1083
QY 61 caagccttgaagagagatgtccctctgtataatgtcaagcagctgtccctccatccatcc 120
DB 1084 CAAGCCTTGAAGAGAGATGTCCCTCTGTATAATGTCAAGCAGCTGTCCCTCTCCGATAC 1143
QY 121 tctcccaactcacctctggagagtgagctgtctcaatccagagctctctgctgc 180
DB 1144 TCTCGCAACTCACGCTCTGGGAGTGGGCTGTCTGCAATCCAGAGCTGTCTGCTGC 1203
QY 181 tcggaccaccagcactgtgccccagcgatacacagtgtgtagctgagggcagtgctcag 240
DB 1204 TCGGACCACCAGCAGTGTGCCCCCAGCATACAGCTGTGTAGCTGAGGGGCAAGTGTACG 1263
QY 241 cgaggaagcagatcgtggtgactggaagatgctgcgcgcgcgcgcgcgcgcgcgcgcgc 300
DB 1264 CGAGGAAGCGAGATCGTGTGCTGAGTGGAGAGATGCTGTGCCCGCGCGCTTCTTATCC 1323
QY 301 cccccagagacatcgctgtaccagacacaccagctgcccgtggggcgaacctgctgc 360
DB 1324 CACCCAGAGACATCGGCTGTACACAGCACACAGCTGCCCGGTGGCGGGAACCTGTCTGC 1383
QY 361 ccgagccaggtgagagctgggctgctccagttgccccatgctgtgtgctgcagagat 420
DB 1384 CCGAGCCAGGTTGGAGCTGTGGGCTGTCTGCCAGTGTGCCCATGCTGTGTGCTGCCAGGAT 1443
QY 421 cgcagcactgtgcccggctggctacacctgcaactggaagctcgatccctgcgagaag 480
DB 1444 CGCCAGCAGTGTGCCCGCTGGCTGTACACTGCAAGCTGAAGCTCGATCTCGAGAGAG 1503
QY 481 gaagtgtctctccagcctgcccactctctgcccgttagcctcaactcagctgggtggaag 540
DB 1504 GAAGTGTCTCTGCCAGCCTGCCACTCTCTGGCCCGTGAAGCTCACCTGCTGGTGAAG 1563
QY 541 gacgtgagtgagggaagacactctgccatgataaacagacactgctgcgcgagacac 600
DB 1564 GACGTGGAGTGGGGAAGACACTCTGCCATGATTAACAGACTGTCTGCCGAGACAC 1623
QY 601 cgacagggctggcctgtctccatcgccccagggcgctgtgtgtgctgagcgccac 660
DB 1624 CGACAGGCTGGCCCTGTCTCTACGCCAGGGCTGTGTGTGCTGTATCGGCCAC 1683
QY 661 tgcgtcctgctgctcgcctgcgcagcagggggtacaaagtgttgccagggagggc 720
DB 1684 TGCTGTCTGTCTGCTGTCTGCTGCGCACGAGGGGTACCAAGTGTGTGCGCAGGGAGGC 1743
QY 721 ccgcctggagcgcctcttgagggaccagccttgagacagctgtgtgagggacagta 780
DB 1744 CCGCGTGGGAGCGCCCTTTGAGGGACCCAGCCTTGAGACAGCTGTGTGAGGGACAGTA 1803
QY 781 ctgaagactctgagcctcctcgagccccactcggaggggtgcccctgtcagggcctccct 840
DB 1804 CTGAAGACTCTGAGCCCTCGGAGCCCACTCGAGGGGTGCCCTGTCTGCTCAGGGCTCCCT 1863
QY 841 agcactccccctaaacaaattctccctgagaccctattgagctcccatcacatggg 900
DB 1864 AGCAGCTCCCCCTAAACAAATTCTCCCTGGACCCCAATTCTGAGCTCCCATCACCATGGG 1923
QY 901 agtggggctcaatctaaggcccttccctgtcagagggggttgaggcaaaagccatt 960
DB 1924 AGTGGGGCTCAATCTAAGGCCCTTCCCTGTCTCAGAGGGGGTTGAGGCAAAAGCCCAAT 1983
QY 961 acaagctgcacccccctccctgttcagtgagccctgtgcccaggtgtctttccctatcc 1020
DB 1984 ACAAGCTGCACATCCCTCCCGCTTTCAGTGGACCCCTGTGCGCAGGTGTCTTCCCTATCC 2043
QY 1021 acagggatgttgtgtgaggtgtctttcaataaagtgttcactttctt 1072
DB 2044 ACAGGGGTGTGTGTGTGGGTGTCTTCAATAAAGTTGTGCACTTCTT 2095
```

RESULT 2  
HUMGRANUL  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
FEATURES  
source

HUMGRANUL  
H.sapiens granulin mRNA, complete cds.  
M75161  
M75161.1 GI:183612  
granulin.  
Homo sapiens (tissue library: Clontech HL1058b) adult bone marrow mRNA.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Bhandari,V., Palfree,R.G. and Bateman,A.  
Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains  
Proc. Natl. Acad. Sci. U.S.A. 89 (5), 1715-1719 (1992)  
92179253  
Location/Qualifiers  
1..2095  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="bone marrow"  
/dev\_stage="adult"  
/tissue\_lib="Clontech HL1058b"  
1..2095  
/gene="granulin"  
13..1734  
/gene="granulin"  
/codon\_start=1  
/product="granulin"  
/protein\_id="AAA58617.1"  
/db\_xref="GI:183613"  
/translation="MTLVSWALTAGLVAGTRCPDGFVACCLDPGASVSCCRP  
LDRWPTLSRLHGGPCQVDAHCSAGHSCEFTVSGTSSCPPEAVACGDGHHCPRG  
FHCSDARSQFSGNSVGAIQCDPSQDFCTCCVMDGSMGCPMPQASCCD  
RVHCCPHGAFCDLVHTRCITPTGTHPLKLPQATNRNALSSVMCPDARSRCPD  
STCCBLPSKYGCCPMPNATCCSHLHCCPQDVTCDLIQSKLSKNATDILLKLP  
HTGVKCDMEVSCPDGTCRQSGAMGCCPFTQAVCCEDHHCPCAGTCDTQGT  
CEQGHQVWHEKAPAHLSLEDPOALKRDVPCDNYSSCPSSDCCQLTSGEWGCP  
EAVCCSDHQCPQRTCVAGCQCGSEIVAGLEKMPARSGLSHPRDIGCDQHTSC  
PVGGTCCPSQSGSWACCOLPHAVCCEDRHCPCAGTTCNVKARSCEKEVVSQA  
PFTL  
ARSPHVGVKDECGEGHFCNDQTCRDNROGWACCPYAQGVCCADRRCPCGPRCA  
RRGTCLRRAPNDALRDPALROLL"  
13..63  
/gene="granulin"  
/note="presumed signal peptide"  
64..1791  
/gene="granulin"  
/note="presumed mature peptide"  
/product="granulin"  
2075..2080  
/gene="granulin"  
BASE COUNT 364 a 682 c 624 g 425 t  
ORIGIN

Query Match 100.0%; Score 1072; DB 9; Length 2095;  
Best Local Similarity 100.0%; Pred. No. 2.5e-234; Indels 0; Gaps 0;  
Matches 1072; Conservative 0; Mismatches 0;  
QY 1 gggccccaccagtgccctggatggagagggcccccagctcacctcagctgagccagaccca 60  
DB 1024 GGGCCCCCACCAGTGCCCTGGATGGAGAGGGCCCCAGCTCACCTCAGCTGCCAGACCCA 1083  
QY 61 caagccttgaagagagatgtccctctgtataatgtcaagcagctgtccctccatccatcc 120  
DB 1084 CAAGCCTTGAAGAGAGATGTCCCTCTGTATAATGTCAAGCAGCTGTCCCTCTCCGATAC 1143  
QY 121 tctcccaactcacctctggagagtgagctgtctcaatccagagctctctgctgc 180  
DB 1144 TCTCGCAACTCACGCTCTGGGAGTGGGCTGTCTGCAATCCAGAGCTGTCTGCTGC 1203  
QY 181 tcggaccaccagcactgtgccccagcgatacacagtgtgtagctgagggcagtgctcag 240  
DB 1204 TCGGACCACCAGCAGTGTGCCCCCAGCATACAGCTGTGTAGCTGAGGGGCAAGTGTACG 1263  
QY 241 cgaggaagcagatcgtggtgactggaagatgctgcgcgcgcgcgcgcgcgcgcgcgcgc 300  
DB 1264 CGAGGAAGCGAGATCGTGTGCTGAGTGGAGAGATGCTGTGCCCGCGCGCTTCTTATCC 1323  
QY 301 cccccagagacatcgctgtaccagacacaccagctgcccgtggggcgaacctgctgc 360  
DB 1324 CACCCAGAGACATCGGCTGTACACAGCACACAGCTGCCCGGTGGCGGGAACCTGTCTGC 1383  
QY 361 ccgagccaggtgagagctgggctgctccagttgccccatgctgtgtgctgcagagat 420  
DB 1384 CCGAGCCAGGTTGGAGCTGTGGGCTGTCTGCCAGTGTGCCCATGCTGTGTGCTGCCAGGAT 1443  
QY 421 cgcagcactgtgcccggctggctacacctgcaactggaagctcgatccctgcgagaag 480  
DB 1444 CGCCAGCAGTGTGCCCGCTGGCTGTACACTGCAAGCTGAAGCTCGATCTCGAGAGAG 1503  
QY 481 gaagtgtctctccagcctgcccactctctgcccgttagcctcaactcagctgggtggaag 540  
DB 1504 GAAGTGTCTCTGCCAGCCTGCCACTCTCTGGCCCGTGAAGCTCACCTGCTGGTGAAG 1563  
QY 541 gacgtgagtgagggaagacactctgccatgataaacagacactgctgcgcgagacac 600  
DB 1564 GACGTGGAGTGGGGAAGACACTCTGCCATGATTAACAGACTGTCTGCCGAGACAC 1623  
QY 601 cgacagggctggcctgtctccatcgccccagggcgctgtgtgtgctgagcgccac 660  
DB 1624 CGACAGGCTGGCCCTGTCTCTACGCCAGGGCTGTGTGTGCTGTATCGGCCAC 1683  
QY 661 tgcgtcctgctgctcgcctgcgcagcagggggtacaaagtgttgccagggagggc 720  
DB 1684 TGCTGTCTGTCTGCTGTCTGCTGCGCACGAGGGGTACCAAGTGTGTGCGCAGGGAGGC 1743  
QY 721 ccgcctggagcgcctcttgagggaccagccttgagacagctgtgtgagggacagta 780  
DB 1744 CCGCGTGGGAGCGCCCTTTGAGGGACCCAGCCTTGAGACAGCTGTGTGAGGGACAGTA 1803  
QY 781 ctgaagactctgagcctcctcgagccccactcggaggggtgcccctgtcagggcctccct 840  
DB 1804 CTGAAGACTCTGAGCCCTCGGAGCCCACTCGAGGGGTGCCCTGTCTGCTCAGGGCTCCCT 1863  
QY 841 agcactccccctaaacaaattctccctgagaccctattgagctcccatcacatggg 900  
DB 1864 AGCAGCTCCCCCTAAACAAATTCTCCCTGGACCCCAATTCTGAGCTCCCATCACCATGGG 1923  
QY 901 agtggggctcaatctaaggcccttccctgtcagagggggttgaggcaaaagccatt 960  
DB 1924 AGTGGGGCTCAATCTAAGGCCCTTCCCTGTCTCAGAGGGGGTTGAGGCAAAAGCCCAAT 1983  
QY 961 acaagctgcacccccctccctgttcagtgagccctgtgcccaggtgtctttccctatcc 1020  
DB 1984 ACAAGCTGCACATCCCTCCCGCTTTCAGTGGACCCCTGTGCGCAGGTGTCTTCCCTATCC 2043  
QY 1021 acagggatgttgtgtgaggtgtctttcaataaagtgttcactttctt 1072  
DB 2044 ACAGGGGTGTGTGTGTGGGTGTCTTCAATAAAGTTGTGCACTTCTT 2095

Db 1144 TGCTGCCAATCACTGCTGGGAGTGGGGCTGCTGTGTCCTCAATCCAGAGGCTGTCTGCTGC 1203  
Qy 181 tcgaccaccagcactgtctccccccagcagatcacagctgtgttagctgagggcgagtgctcag 240  
Db 1204 TCGGACCACAGCACTGCTGCCGCCAGCATACAGCTGTGTAGCTGAGGGCGAGCTGCAG 1263  
Qy 241 cgaggagcagagatcgtggtgagctgagagagatgctgcccgcgcggttccattatcc 300  
Db 1264 CGAGGAACGAGATCGTGGCTGAGCTGAGAGATGCTGCCGCGCGGTTCCTTATCC 1323  
Qy 301 caccaccagacatcgtctgaccagcacaccagctgcccgtggcgagcagcctgctgc 360  
Db 1324 CACCCAGAGACATCGGCTGTGACACACACAGCTGCCCGGTGGCGGAACCTGCTGC 1383  
Qy 361 ccgagcagagtggtgagctggtgctgctgagctgcccagctgctgtgtgctgcagagat 420  
Db 1384 CCGAGCCAGGCTGGGAGCTGGGCTGCTGCCAGCTGCTGCCCATGCTGTGTCTGGAGAT 1443  
Qy 421 cgcagcactgctgcccggctggtgctacacctgaacgtgaagctgagctgagctgcgagag 480  
Db 1444 CGCCAGCACTGCTGCCCGGCTGCTACACCTGCAACGTGAAGGCTGATCTGCGAGAG 1503  
Qy 481 gaagtgtctcgcagcagctgcaacctctcgtgcccagctgagcctgagcctcagctggtgag 540  
Db 1504 GAAGTGTCTCTGCCAGCTGCTGCCACCTTCTGTCGCGGTGAGCTGCTGCTGCTGAG 1563  
Qy 541 gacgtgagtggtggagacacttctgcatgataaacagacactgctgcccagagacaac 600  
Db 1564 GACGTGAGTGTGGGAGGACACTTCTGCCATGATTAACAGACTGCTGCCGAGAGAC 1623  
Qy 601 cgacagggctgggctgctgtccctacgcccagggcgctgtgtgtgctgagctgcgcgac 660  
Db 1624 CGACAGGCTGGGCTGCTGTGCTGCTAGCCGCGGCTGCTGTGCTGATGCTGCGGCGAC 1683  
Qy 661 tgcgtcctgctgctcctgctgagcagcaggggtaccagtgctgtgagggagggcc 720  
Db 1684 TGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1743  
Qy 721 ccgcgtggagcgcctcttgaggagccagccttgagagagctgctgtgagggagagta 780  
Db 1744 CCGGCTGGGAGCGCCCTTGGAGGAGCCAGCTGAGACAGCTGCTGAGGAGACAGTA 1803  
Qy 781 ctgaagactctgagcctcctggagccacactcggaggggtgctctgctgagggcctcct 840  
Db 1804 CTGAAGACTCTGAGCCCTCGGAGCCACCTGAGGAGGCTGCTGCTGCTGCTGCTGCT 1863  
Qy 841 agcactcctcctaaccaaatctcctgagcaccctattgagctcctcctcctcctcctcct 900  
Db 1864 AGCACTCCCTTAACCAATTCCTGCTGAGCCCTTCTGAGCTGCTGCTGCTGCTGCT 1923  
Qy 901 aggtgggctcactcctaaagcctcctcctgagaggggtgagggagggagggagggaggg 960  
Db 1924 AGTGGGGCTCAATCTAAGCCCTTCTGCTGAGAGGGGTTGAGGCAAAAGGCCAT 1983  
Qy 961 acaagctgccatccccctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1020  
Db 1984 ACAAGCTGCCATCCCTCCCTGCTTCAAGTGGAGCTGCTGCTGCTGCTGCTGCTGCT 2043  
Qy 1021 acaggggt 1072  
Db 2044 ACAGGGGT 2095

RESULT 3  
BC010577 BC010577 2132 bp mRNA linear PRI 12-JUL-2001  
LOCUS Homo sapiens, granulin, clone MGC:9342 IMAGE:3457813, mRNA,  
DEFINITION complete cds.  
ACCESSION BC010577  
VERSION BC010577.1 GI:14714850  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2132)  
Strausberg, R.  
Direct Submission  
Submitted (10-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: AFCC  
CDNA Library Preparation: Life technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: villalona@bcm.tmc.edu.  
Villalona, D.K., Luna, R.A., Hale, S.M., Rulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAK Plate: 12 Row: a Column: 17  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 10435243.  
FEATURES  
Location/Qualifiers  
1. .2132  
/organism="Homo sapiens"  
/db\_xref="LocusID:2896"  
/db\_xref="taxon:9606"  
/clone="MGC:9342 IMAGE:3457813"  
/tissue\_type="Cervix, carcinoma"  
/clone\_lib="NIH-MGC\_12"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
32. .1813  
/codon\_start=1  
/product="granulin"  
/protein\_id="AAH10577.1"  
/db\_xref="GI:14714851"  
/translation="MTLTVSWALTAGLVAGTRCPDQGFVACCLDPGGASYSCCRP  
LLDKWPTLSRLHGGPCQVDHAGSAGHSCTFTVSTSCCPPEAVAGDGHGCCPRG  
FHCSDGRSCFORSGNNSVGAIOCPDQFECDFSTCCVMVDGSCCPMPQASCCED  
RVCCCPHAGCDLVHTCITPTGTHPLAKLPAQRTNRAVALSSVMCPDARSRCPDG  
STCELPFGKTCGCCPMFNATCCSHLHCCPDQDVTCDLIQSKLKENATDILLKPLA  
HTVGDVKECMVECPDGYTCCLRGSGAWGCCPFTQAVCCEDHICCPAGFTCDTQKGT  
CEGPHQVPMWEKAPAHLSLPDQALRDPVDNVSPPSSDTCCLTSGGWCPIP  
EAVCCSDHCCPCPGYTCVAGCGQCGSEIVAGLEKMPARRASLSHPDIDCDHTSC  
PVGTCPCPSLGGSWACCOLPHAVCCEDHCCCPAGYTCNVKARCEKEVSAQATFL  
ARSPHGVKDVCEGEGHCFHDNPTCCRDNRQGWACCPYRQGVCCADRRHCCPAGFRCA  
ARSTKLLRREAPRDAPDRPALAKLL"  
BASE COUNT 383 a 687 c 636 g 426 t  
ORIGIN

Query Match 92.6%; Score 992.6; DB 9; Length 2132;  
Best Local Similarity 98.5%; Pred. No. 3.3e-216;  
Matches 1023; Conservative 0; Mismatches 14; Indels 2; Gaps 2;  
Qy 1 gggccaccaggtgctgagtgagagggccagctcactcagctgagcagccca 60  
Db 1043 GGGCCCCACAGGTGCTGTGATGAGAGGCCCGCCAGCTCAGCTGCGAGACCCA 1102  
Qy 61 caagccttgaagagatgtccctgtgataatgtcagcagctgtccctcccgatacc 120  
Db 1103 CAAGCCTTGAAGAGAGATGTCCTCTGTATATGTGAGCAGCTGCTCCCTCCGATACC 1162  
Qy 121 tgcgtccaaactcagctctgaggagtgagggtggtggtggtggtggtggtggtggtggtg 180



```
QY 1 gggccccacagtgccctggatgaggaagggccccagctcaactcagcctgcccagacc 60
Db 1052 GGGCCCCACAGTGCCTGGATGAGAGAGGCCAGCCAGCTCACCTCAGCTGCCAGACCA 1111
QY 61 caagccttaagagagatccccctgataatgtcagcagctgtccctccctccgatacc 120
Db 1112 CAAGCCTTGAAGAGAGATGCCCTGTGATATGTGAGAGCTGTCCCTCCCGATACC 1171
QY 121 tgcgtccaaactcagctctggggagtggtgctgtgtccaaatccagagctgtctgtgc 180
Db 1172 TGCTGCCAACTCAGCTGTGGGAGTGGGCTGTGTGTCATCCAGAGCTGTCTGTCTGC 1231
QY 181 tcggaccacagcactgtgtccccagcgatacacagtgtgtagctgaggggagtgatcag 240
Db 1232 TCGGACCACAGCAGCTGTGCCGCCAGGCTACACCTGTGTAGCTGAGGGGAGTGTCTAG 1291
QY 241 caggaagcagagatcgtgctgagctgaggaagatccctgccgcgcggttcttacc 300
Db 1292 CGAGGAAGCAGAGATCGTGGCTGGACTGGAGAGATGCTGCCCGCGGGCTTCTTATCC 1351
QY 301 caccacagacatcgctgtgaccagcacacagctgcccgtgtggcggaacctgtgc 360
Db 1352 CACCCAGAGACATGGCTGTGACAGCAGACACAGCTGCCGTGGGGGAGAGCTGCTGC 1411
QY 361 ccgagccaggtgagctgggctgctgcccagctgcccagctgtgtgtgctgcgagat 420
Db 1412 CCGAGCCTGGTGGAGCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1471
QY 421 cgcagcactgtgcccgggtgctacacactgcaacgtgaagctcgatcctcgagaag 480
Db 1472 CGCAGAGCTGCTGCCGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1531
QY 481 gaagtgtctctccagcctgcacactctcgtccgtgagcctcagctgagtgagag 540
Db 1532 GAAGTGTCTCTGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1591
QY 541 gacgtgagtggtgggaagacacttctgcatgataaacagacactgctccgagacaac 600
Db 1592 GACGTGAGTGTGGGAAGGACACTTCTGCATGATTAACAGACCTGCTGCCGAGACAAC 1651
QY 601 cgcagggctgggctgctgctcctcagccagggcgctgtgtgtgtgtgtgtgtgtgt 660
Db 1652 CGACAGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
QY 661 tgcgtcctgctgctcctcctgctgctgctgctgctgctgctgctgctgctgctgct 720
Db 1712 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1771
QY 721 ccgcgtggagacgcccctttgagggaccacagccttgagacagctgtgtgtgaggaagta 780
Db 1772 CCGGCTGGGAGCGCCCTTTGAGGAGACCCAGCCTTGAGACAGCTGCTGCTGAGGACAGTA 1831
QY 781 ctgaagactctgagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 840
Db 1832 CTGAAGACTCTGACAGCCTCGGGACCCACCTCGAGGGGTGCTGCTGCTGCTGCTGCTGCT 1891
QY 841 agcacctccccttaacaaattctcctgacccattctgagctcccacatcattcattgag 900
Db 1892 AGCAGCTTCCCTTAACCAATCTCCTGTGACCCCATCTCTGAGCTTCCCATCAGCATGGG 1951
QY 901 aggtggggcctcaatcctaaagcccttccctgtcagaaaggggttgagggcaaaagcc-cat 959
Db 1952 AGGTGGGGCCTCAATCTAAGG-CCTTCCCTGTGAGAGGGGGTGTGGGCAAAAGCCACAT 2010
QY 960 tacaagctgcaatccctccctccctccctccctccctccctccctccctccctccctccct 1019
Db 2011 TACAAGCTGCAATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2070
QY 1020 cacaggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1038
Db 2071 CACAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2089
```

## RESULT 5

```
AF055008
LOCUS Homo sapiens clone 24720 epithelin 1 and 2 mRNA, complete cds.
DEFINITION Homo sapiens clone 24720 epithelin 1 and 2 mRNA, complete cds.
ACCESSION AF055008
VERSION AF055008.1 GI:3005729
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2178)
AUTHORS Yu, W., Andersson, B., Worley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
TITLE Ricafrente, J.Y., Wentland, M.A., Lennop, G. and Gibbs, R.A.
JOURNAL A "double adaptor" method for improved shotgun library construction
MEDLINE Anal. Biochem. 236 (1), 107-113 (1996)
96207227
REFERENCE 2 (bases 1 to 2178)
AUTHORS Yu, W., Andersson, B., Worley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
TITLE Ricafrente, J.Y., Wentland, M.A., Lennop, G. and Gibbs, R.A.
JOURNAL Large-scale concatenation cDNA sequencing
MEDLINE Genome Res. 7 (4), 353-358 (1997)
97264341
REFERENCE 3 (bases 1 to 2178)
AUTHORS Yu, W. and Gibbs, R.A.
JOURNAL Direct Submission
MEDLINE Submitted (23-MAR-1998) Molecular and Human Genetics, Baylor
TITLE College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
JOURNAL Location/Qualifiers
FEATURES
source
1..2178
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE Consortium 24720"
/sex="female"
/tissue_type="brain"
/clone_lib="INIB"
/dev_stage="infant"
63..1844
/note="similar to Homo sapiens epithelin 1 and 2 encoded
by GenBank Accession Number X62320"
/codon_start=1
/product="epithelin 1 and 2"
/protein_id="AAC09359.1"
/db_xref="GI:3005730"
/translation="MWTLVSVALTAGLVAGTRCPDQGFQCPVACCLDPGASYSCCRPL
LDKWPITLRLHGLGCPQVDAHSCAGHSCLFTVSGSCFPFPEAVACGPHGCCPRG
FHCSDAGRSFCQSRNNNSVGAIQCPDQFQCPDFSTCCVMVDGSGCCPMPASCCED
RVHCCPHGAFCDLVHTRCTIPTGTHPLAKLPAORTNRAVLSVSMCPDARSRCPDG
STCCCLPSCKYGCCPNATCCSDHLHCCPQDTCVCLIQSKLSKENATVLLTKLPA
HTVGDVDCMEVSKCPDGYCCRLQSGANGCCPFTQAVCCEDHILHCCPAGTCPTQKGT
CEGPHQVPMMEKAPAHLSLPDPAKLRDVPDNDVSCPSDSDTCCQLTSGEWCCPPI
EAVGCSHDHCCPQGYTCVAEGCQQRGSEIVAGLEKMPARRASLSHPRDITGCDHTSC
PVGTCPCPSLGGSWACQQLPHAYCCEDRQHCPCAGTCTCNKARSCEKVEVSAQPAFL
ARSPHGVKDVCEGEGHFDHNOTCCRDNRQGWACCPYRGVGVCCADRRHCCPAGFCA
ARSTKCLREAPHPDAPLRDPAIRLL"
BASE COUNT 401 a 697 c 646 g 434 t
ORIGIN
```

```
Query Match 92.6% Score 992.6; DB 9; Length 2178;
Best Local Similarity 98.5%; Pred. No. 3.3e-216;
Matches 1023; Conservative 0; Mismatches 14; Indels 2; Gaps 2;
QY 1 gggccccacagtgccctggatgaggaagggccccagctcaactcagcctgcccagacc 60
Db 1074 GGGCCCCACAGTGCCTGGATGAGAGAGGCCAGCCAGCTCACCTCAGCTGCCAGACCA 1133
QY 61 caagccttaagagagatgtccctgtgataatgtcagcagctgtccctccctcagatacc 120
Db 1134 CAAGCCTTGAAGAGAGATGTCCCTGTGATATGTGAGAGCTGTCCCTCCCGATACC 1193
QY 121 tgcgtccaaactcagctctggggagtggtgctgtgtccaaatccagagctgtctgtgc 180
Db 1194 TGCTGCCAACTCAGCTGTGGGGAGTGGGGCTGTGTGTCATCCAGAGCTGTCTGTCTGC 1253
```













```
Qy 529 gtgggtggaagcagctgagtgagggaagacacattctcccatgataacacacacgc 588
Db 3309 GTGGGTGGAAGAGCTGGAGTGGGGAAGACACTTCTCCATGATACACACCTGC 3368
Qy 589 tgcgagacaacacagcagggctggcctgctgtccctacgc----- 629
Db 3369 TGCCGAGACAACCGACAGGGCTGGGCTGCTGCTCCACGCCAGGTGAGCAACNCA 3428
Qy 630 ----- 629
Db 3429 TCCTGGGGNTGATGCGCAGGACACAGTCCNACCTNTGTCACACCTCTCGCNNNNCTCT 3488
Qy 630 -----ccagggctgtgtgctgacgcgcacgtgctgctgctgctgctgctgctg 683
Db 3489 GACCATCCAGGGCTGCTGTGTGATGCGGCGCACCTGCTGCTGCTGCTGCTGCTG 3548
Qy 684 cgcagcaggggtaccagtgattgcagggagggccgcgcgtgggagcgccttttgag 743
Db 3549 CGCAGCCAGGGGTACCAAGTCTTTGCGCAGGAGGCCCGCGCTGGGACGCCCTTTGAG 3608
Qy 744 gacccagccttgacagctgc-tgtgaggacagtagtaagaactctgagccctcgg 802
Db 3609 GGACCCAGCCTTGACAGAGCTGCTGTGTGAGGAGACAGTACTGAAGACTCTGCAGCCCTCGG 3668
Qy 803 gacccactcggagggtgcccctgctcagcgcctccctcagcctcccccctaaacaaatt 862
Db 3669 GACCCCACTCGAGGGTGCCCTCTGCTCAGGCGCCCTCCCTAGCACCTCCCTTAACCAATT 3728
Qy 863 ctccctgagccccattctgagctcccccaccatcaccatggaggtggggtcctaatcattgagc 922
Db 3729 CTCCTCGACCCCTTCCTGAGCTCCCATCACCATGCGAGGTGGGCGCTCAATCTAAGGC 3788
Qy 923 ccttcctctcagagggggtgaggcaaaagcc-cattacaagctgcacccctccccc 981
Db 3789 CTTTCCCTGTGAGAGGGGGTGTGGCAAAAGCCACATTAGAAGCTGCATCCCTCCCTCC 3848
Qy 982 gtttcagtgagccctgtgagcaggtgctttccctatccacaggggtgtgtgtgt 1038
Db 3849 GTTTCAGTGGACCTGTGGCCAGGTGCTTTCCCTATCCACAGGGGTGTGTGTGTGT 3905

RESULT 12
AC019152 Homo sapiens 180520 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 17 clone RP11-812N9, WORKING DRAFT
DEFINITION SEQUENCE, 12 unordered pieces.
AC019152
VERSION AC019152.5 GI:9570407
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
WATERSTON R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 180520)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jun 17, 2000 this sequence version replaced gi:7524009.
```

```
Sequencing vector: M13; 83%
Chemistry: Dye-primer ET; 83% of reads
Chemistry: Dye-terminator Big Dye; 17% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 174306 bases at least Q40
Consensus quality: 176607 bases at least Q30
Consensus quality: 177668 bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 179420; sum-of-contigs
Quality coverage: 7.04 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1. 1099: contig of 1099 bp in length
* 1100: gap of unknown length
* 1200: contig of 1269 bp in length
* 2468: gap of unknown length
* 2569: contig of 1579 bp in length
* 4147: gap of unknown length
* 4148: gap of unknown length
* 4248: contig of 1092 bp in length
* 5340: gap of unknown length
* 5440: contig of 1058 bp in length
* 6498: gap of unknown length
* 6598: contig of 1577 bp in length
* 8175: gap of unknown length
* 8275: contig of 6297 bp in length
* 14571: gap of unknown length
* 14671: contig of 10095 bp in length
* 24767: gap of unknown length
* 24867: contig of 14273 bp in length
* 39140: gap of unknown length
* 39240: contig of 23562 bp in length
* 62802: gap of unknown length
* 62902: contig of 50553 bp in length
* 113454: gap of unknown length
* 113555: contig of 66966 bp in length.
FEATURES
Location/Qualifiers
1..180520
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-812N9"
1..1099
/misc_feature
1200..2468
/misc_feature
2569..4147
/misc_feature
4248..5339
/misc_feature
5440..6497
/misc_feature
6598..8174
/misc_feature
8275..14571
/misc_feature
14672..24766
/misc_feature
24867..39139
/misc_feature
39240..62801
/misc_feature
62902..113454
/misc_feature
```





[illegible]



Db 1791 AACGACTAAAGAACTCCACAGTCTCTGGGAACCCCTGTTCTGAGGGTATCCACCACCTCAGGC 1850  
QY 835 CTCCTAGCACCTCCCTCCCTAACCAAAATCTCCCTGGACCCCATCTGAG-CTCCCATCA 893  
Db 1851 CTCCTGGCACCTCTTCCTTTT--AGTCTCCCGGCCCTACTCATCTGAGTCACCCCATCA 1908  
QY 894 CCATGGAGGTGGGGCCTCAAACTAAGACCTTCTTTATGGAAGAGGCTGTGGCAAAA 1968  
Db 1909 CCATGGAGGTGGGGCCTCAAACTAAGACCTTCTTTATGGAAGAGGCTGTGGCAAAA 1968  
QY 954 GCCATTACAAGTGCATCCCTCCCTCCCTGTTCAAGTGACCCCTGTGGCAAGGTGCTTTTC 1013  
Db 1969 GCCCGTATCAAACTGCCATTTCTATGATTTCTGTGGACCTTGTGGCCAGGTGCTCTTC 2028  
QY 1014 CCTATCACAGGTGTTGTGTGTTGTTGTTG 1046  
Db 2029 TCGATCCACAGGTGTTGTGTGAGCTTCTTTTG 2061

Search completed: September 5, 2002, 16:33:29  
Job time: 12132 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2002, 18:30:02 ; Search time 457.59 Seconds  
(without alignments)  
4022.230 Million cell updates/sec

Title: us-09-824-647-16\_COPY\_1024\_2095  
Performance: 1072  
Sequence: 1 gggccccaccagtgccctg.....ataagttgtgacttttt 1072

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

- 1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1070.4	99.9	2095	20 AAV82825	Human GP88 autocri
2	1070.4	99.9	2124	14 AAQ56794	Granulin coding se
3	1070.4	99.9	2198	14 AAQ49052	Granulin coding se
4	992.6	92.6	2432	21 AACT78180	Human cancer assoc
5	989.4	92.3	1630	22 AAH16370	Human cDNA sequenc
6	747.2	69.7	1779	12 AAQ14339	Human epithelin pr
7	535.4	49.9	2137	20 AAQ62824	Mouse GP88 autocri
8	483.8	45.1	561	21 AAA69779	Human ovarian carc
9	469.2	43.8	1767	12 AAQ14338	Rat epithelin prec

10	448.4	41.8	1767	12	AAQ14340	Mouse epithelin pr
11	339.2	31.6	439	16	AAQ22115	Human gene signatu
12	228.8	21.3	565	22	AAK18681	Human brain expres
13	206.6	19.3	565	22	AAK53492	Murine transport a
14	197.8	18.5	336	22	AAK54602	Human ovarian PCR-
15	197.8	18.5	336	22	AAH83224	Human ovarian tumo
16	195.6	18.2	390	22	AAK05895	Human brain expres
17	153.2	14.3	539	12	AAQ14952	Bovine epithelin p
18	142.2	13.3	757	22	AAH07141	Human cDNA clone (
19	65.4	6.1	430	22	AAI82981	Human polynucleoti
20	64.6	6.0	341	12	AAQ14953	Chicken epithelin
21	57	5.3	177	22	AAI15045	Oryctolagus cunicu
22	51.8	4.8	561	21	AAK69780	Human ovarian carc
23	43.8	4.1	561	21	AAK69780	Human ovarian carc
24	43.4	4.0	1806	21	AAK50102	Human Znt2 consen
25	43	4.0	2608	22	AAK98385	Human cDNA clone B
26	38.2	3.6	199	14	AAQ49051	Probe (grnA) used
27	38.2	3.6	114955	20	AAK53491	Human adenosine A1
28	38	3.5	774	17	AAI13276	DNA encoding uropo
29	38	3.5	6741	21	AAK10595	Gene encoding a su
30	37.8	3.5	10732	21	AAK10594	Gene encoding a su
31	37	3.5	20303	18	AAK71699	Human deoxycytidy
32	37	3.5	26764	18	AAK71696	Human deoxycytidy
33	36.6	3.4	461	22	AAK93033	Human cDNA 3'-end
34	36.6	3.4	1721	24	AAK09743	Human ovarian tumo
35	36.6	3.4	2348	22	AAH17671	Human cDNA sequenc
36	36.4	3.4	1501	22	AAH24640	Human PEG1 DNA. B
37	36.4	3.4	2497	22	AAH14473	Human cDNA sequenc
38	36.4	3.4	3789	23	AAK86522	DNA encoding novel
39	36.4	3.4	18678	22	AAK32889	Human genomic DNA
40	36.4	3.4	18678	22	AAK91271	Human digestive sy
41	36.2	3.4	4403765	22	AAI99683	Mycobacterium tube
42	36	3.4	935	21	AAK29743	Human gamma 1 heav
43	36	3.4	935	21	AAK39881	Human gamma 1 heav
44	36	3.4	935	22	AAK35599	Human gamma 1 heav
45	36	3.4	1383	21	AAK29742	Human gamma 1 heav

ALIGNMENTS

RESULT 1

AAV82825	ID	AAV82825 standard; cDNA; 2095 BP.
XX	XX	
AC	AAV82825;	
XX	XX	
DT	15-MAR-1999 (first entry)	
XX	XX	
DE	Human GP88 autocrine growth factor cDNA.	
XX	XX	
KW	GP88; granulin; epithelin; human; growth factor; autocrine; tumour;	
KW	cancer; viral infection; antagonist; therapy; diagnosis; ss.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
FH	Key	Location/Qualifiers
FT	CDS	13..1794
FT		/*tag= a
XX	XX	
PN	W09852607-A1.	
XX	XX	
PD	26-NOV-1998.	
XX	XX	
PF	22-MAY-1998; 98WO-US10555.	
XX	XX	
PR	16-DEC-1997; 97US-0991862.	
PR	23-MAY-1997; 97US-0863079.	
XX	XX	
PA	(SERR/) SERRERO G.	
XX	XX	
PI	Serrero G;	
XX	XX	



**Oy**    1   gggcccaacagtgccctgtagtggagaaggcccagctcaccctcagctgccagacca   60  
**Db**    1111   gggcccaacagtgccctgtagtggagaaggcccagctcaccctcagctgccagacca   1170	
**Oy**    61   caagccttgaagagagatgccctgtgataatgtcacgagtgtccccctcccgatacc   120

08-FEB-2001 (first entry)

Human cancer associated gene sequence SEQ ID NO:574.

Human: cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral; antinflammatory; antithyroid; antiallergic; antibacterial; cardiant; dermatologic; neuroprotective; thrombolytic; coagulant; nootropic; vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation; immune disorder; haematopoietic cell disorder; autoimmune disorder; allergic reaction; graft versus host disease; organ rejection; haemostatic; thrombolytic; cardiovascular disorder; infection; neurological disease; drug screening; ss.

Homo sapiens.

W0200055350-A1.

21-SEP-2000.

08-MAR-2000; 2000WO-US05882.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-587533/55.

P-PSDE; AAB43971.

Novel isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer -

Claim 1; Page 1097-1098; 2352pp; English.

AAC77607 to AAC78448 encode the human cancer associated proteins given in AAB43398 to AAB44239. The proteins can have activities based on the tissues and cells the genes are expressed in. Example of activities include: cytostatic; proliferative; vulnerary; immunomodulator; antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antinflammatory; antithyroid; antiallergic; antibacterial; antiviral; dermatologic; neuroprotective; cardiant; thrombolytic; coagulant; polynucleotides and polypeptides can be used for preventing, treating or ameliorating medical conditions and diagnosing pathological conditions. Polynucleotides, polypeptides, antibodies, agonists and antagonists from the present invention may be used to treat immune disorders by activating or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haematopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostatic or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological disease and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may be also be used in drug screens. AAC78449 to AAC78457 and AAB44240 represent sequences used in the exemplification of the present invention.

Sequence 2432 BP; 460 A; 764 C; 707 G; 496 T; 5 other;

Query Match 92.6%; Score 992.6; DB 21; Length 2432;  
Best Local Similarity 98.5%; Pred. No. 1.1e-249;  
Matches 1023; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

Qy 1 gggccccaccaggctgccccggatggagagagcccccagctccacctcagctctgcccagatcca 60

Db 1259 caacccctaaagacagatattccccctatgaataatcaccacactatccctccctccgatacc 1318

Db 1199 gggccccaccaggctgcctggatggagaaagcccgagctcacctcagctgcagatcca 1258

Qy 61 caagccttgaagagagatgtccctctgtataatgtcagcagctgtccctccctccgatacc 120

QY	121	tgtgtccaaactcacgtctcgtgggagtggggctgtcgttccaatcccaaggagctgtctctgc.	180
Db			
QY	1319	tgtgtccaaactcacgtctcgtgggagtggggctgtcgttccaatcccaaggagctgtctctgc.	1378
Db			
QY	181	tcggaccaccagcactgtgcgcccaagcatcacagtgttagctgaggcgagctgcag	240
Db			
QY	1379	tcggaccaccagcactgtgcgcccaagctcacacgtgttagctgaggcgagctgcag	1438
Db			
QY	241	csgaggaagcgagatcgtggctggactggagagaatgcctgcgcgcgcggttctcttatcc	300
Db			
QY	1439	csgaggaagcgagatcgtggctggactggagagaatgcctgcgcgcgcggttctcttatcc	1498
Db			
QY	301	cacccacagagacatcgggtgtgaccagcacaccagctgcccggtgggcggaacctctgc	360
Db			
QY	1499	cacccacagagacatcgggtgtgaccagcacaccagctgcccggtgggcgagacctctgc	1558
Db			
QY	361	ccgagccaggggtgggagctgggctgtgcacagttgcccatgtgtgtctgcgaggat	420
Db			
QY	1559	ccgagctgggtgggagctgggctgtgcacagttgcccatgtgtgtctgcgaggat	1618
Db			
QY	421	cgcacgactgtgcgcggctgtcacacctgcaacctggaaggtcgatcctcgcgagaag	480
Db			
QY	1619	cgcacgactgtgcgcggctgtcacacctgcaacctggaaggtcgatcctcgcgagaag	1678
Db			
QY	481	gaagtgtctctgcacagcctgcacaccttctgtgccogtagccctcacgtgggtgtgaag	540
Db			
QY	1679	gaagtgtctctgcacagcctgcacaccttctgtgccogtagccctcacgtgggtgtgaag	1738
Db			
QY	541	gaagtggagtggtgggaagacacttctgccatgataaccagacctgtcccgagacaac	600
Db			
QY	1739	gaagtggagtggtgggaagacacttctgccatgataaccagacctgtcccgagacaac	1798
Db			
QY	601	cgcagggctgggctgtcctcacgcgccagggcgctgtgtgtgatcggcgccac	660
Db			
QY	1799	cgcagggctgggctgtcctcacgcgccagggcgctgtgtgtgatcggcgccac	1858
Db			
QY	661	tgtgtctcgtcgtgtccgctgcgcacgcaggggtaccagtgtttgcgagggaggcc	720
Db			
QY	1859	tgtgtcctcgtcgtgtccgctgcgcacgcaggggtaccagtgtttgcgagggaggcc	1918
Db			
QY	721	ccgcgctgggacgcccctttgaggaccagccttgacacagctgtgaaggacagta	780
Db			
QY	1919	ccgcgctgggacgcccctttgaggaccacgcttgacacagctgtgaaggacagta	1978
Db			
QY	781	ctgaagactctgacgcccctcgggaccccactcggagggtgcctctcgtcacggcctccct	840
Db			
QY	1979	ctgaagactctgacgcccctcgggaccccactcggagggtgcctctcgtcacggcctccct	2038
Db			
QY	841	agcacctccccctaaacaaattctccctggaccccattctgagctccccatcaccatggg	900
Db			
QY	2039	agcacctccccctaaacaaattctccctggaccccattctgagctccccatcaccatggg	2098
Db			
QY	901	aggtggggcctcaatctaaggcccttccctgtcgaagaaggggttgaggcaaaagcc-cat	959
Db			
QY	2099	aggtggggcctcaatctaagg-ccttccctgtcgaagaaggggttgaggcaaaagccaat	2157
Db			
QY	960	tacaagctgcacccccctcccgttttcagtgacccttggccaggtgcttttccctatc	1019
Db			
QY	2158	tacaagctgcacccccctcccgttttcagtgacccttggccaggtgcttttccctatc	2217
Db			
QY	1020	cacaggggtgtttgtgtgt 1038	
Db			
-Db	2218	cacaggggtgtttgtgtgt 2236	
RESULT	5		
AAH16370			
ID	AAH16370 standard; cDNA; 1630 BP.		
XX			
AC	AAH16370;		
XX			
DT	26-JUN-2001 (first entry)		

Human cDNA sequence SEQ ID NO:15309.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J; Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T; WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8; SEQ ID 15309; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 1630 BP; 285 A; 517 C; 489 G; 339 T; 0 other:

Query Match	92.3%	Score	989.4;	DB	22;	Length	1630;
Best Local Similarity	98.3%;	Pred. No.	6.9e-249;				
Matches 1021; Conservative	0;	Mismatches	16;	Indels	2;	Gaps	2;

  

y	1	ggggcccccacaggtgccttggatgagaagggccacagctcaacctcagcctgcagaccaca	60
b	556	ggggcccccacaggtgccttggatgagaagggccacagctcaacctcagcctgcagaccaca	615
y	61	caagccttgaagagagatgtcccttgtataatgtcaagcagctgtccctctcagatacc	120
b	616	caagccttgaagagagatgtcccttgtataatgtcagcagctgtccctctcagatacc	675
y	121	tgctgccaactcaactctctggggagtggggctgctgtccaaatcccaagagctgtctgtgc	180

Db 676 tctgtcccaactcacgtctggggagtggtgctgtgtccatcccaaggctgtctgtgc 735  
Qy 181 tcggaccaccagcactgtgccccagcgatacacgtgtgtagctgagggcagtgctcag 240  
Db 736 tcggaccaccagcactgtgccccagcgctacacgtgtgtagctgagggcagtgctcag 795  
Qy 241 cgaggaagcgagatcgctggactggaggaagatgcctgccccgcgcgttcttattcc 300  
Db 796 cgaggaagcgagatcgctggactggaggaagatgcctgccccgcgcgttcttattcc 855  
Qy 301 caccccagagacatcgctgtgaccagcacacagcactcccggtggcggaacctgtgc 360  
Db 856 caccccagagacatcgctgtgaccagcacacagcactcccggtggcggaacctgtgc 915  
Qy 361 ccgagccaggggtggagctggcctgtgcccagttgccccatgctgtgtgtgcgagat 420  
Db 916 ccgagccaggggtggagctggcctgtgcccagttgccccatgctgtgtgtgcgagat 975  
Qy 421 ccgagccagctgtgccccgctgtgtgtacacctgcaactggaagctcgatcctgcgagaag 480  
Db 976 ccgagccagctgtgccccgctgtgtgtacacctgcaactggaagctcgatcctgcgagaag 1035  
Qy 481 gaagtgtctctgccccagcctgcaacttctgtgccccagcctcagctgaggggtgag 540  
Db 1036 gaagtgtctctgccccagcctgcaacttctgtgccccagcctcagctgaggggtgag 1095  
Qy 541 gacgtggagtggtgggaagacacttgcctatgataccagacactgtcgcgagacaac 600  
Db 1096 gacgtggagtggtgggaagacacttgcctatgataccagacactgtcgcgagacaac 1155  
Qy 601 cgacagggctggcctgtgtccctacgccccagggcgtctgtgtgtgctgacgcccac 660  
Db 1156 cgacagggctggcctgtgtccctacgccccagggcgtctgtgtgtgctgacgcccac 1215  
Qy 661 tctgtctctgt 720  
Db 1216 tctgtctctgt 1275  
Qy 721 ccgctgtggagcgcctcttgaggagaccagccttgagacagctgtgtgtgtgtgtgtgt 780  
Db 1276 ccgctgtggagcgcctcttgaggagaccagccttgagacagctgtgtgtgtgtgtgtgt 1335  
Qy 781 ctgaagactctgagcctctgggacccactcggaggtgcccctgtcagggcctccct 840  
Db 1336 ctgaagactctgagcctctgggacccactcggaggtgcccctgtcagggcctccct 1395  
Qy 841 agcactctccctaaacccaaattctccctggacccttctgaagctcccccacacatgg 900  
Db 1396 agcactctccctaaacccaaattctccctggacccttctgaagctcccccacacatgg 1455  
Qy 901 aggtggggcctcaatctaaggcccttccctgtcagaagggggttgaggcaaaagcc-cat 959  
Db 1456 aggtggggcctcaatctaagg-ccttccctgtcagaagggggttgaggcaaaagcccat 1514  
Qy 960 tacaagctgcaatccctccctccctgttcagtggaacctgtgcccaggtctttccctatc 1019  
Db 1515 tacaagctgcaatccctccctccctgttcagtggaacctgtgcccaggtctttccctatc 1574  
Qy 1020 cacaggggtgtgtgtgt 1038  
Db 1575 cacaggggtgtgtgtgt 1593

## RESULT 6

AAQ14339

ID AAQ14339 standard; DNA; 1779 BP.

XX

AC AAQ14339;

XX

DT 17-JAN-1992 (first entry)

DE

Human epithelin precursor.

XX

KW ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.  
XX Homo sapiens.  
OS Location/Qualifiers  
XX Key 41..1819  
FH misc\_RNA /\*tag= a  
FT /note= \*claim 9, page 54\*  
FT #  
XX  
PN W09115510-A.  
XX  
PD 17-OCT-1991.  
XX  
PF 03-APR-1991; 91WO-US02321.  
XX  
PR 13-MAR-1991; 91US-0083796.  
PR 03-APR-1990; 90US-0504508.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIB.  
XX Shoyab M, Plowman GD;  
XX WPI; 1991-325168/44.  
DR P-PSDB; AAR14326.  
XX  
PT New cysteine-rich growth modulating proteins, epithelins - useful  
PT as inhibitors of neoplastic cell growth and to promote wound  
PT healing and treat psoriasis  
XX  
PS Disclosure; Fig 22; 97pp; English.  
XX  
CC ET-1 and ET-2 were isolated from rat kidneys and their amino acid  
CC sequences determined. A full length rat ET cDNA (AAQ14338) was obtd.  
CC by screening a rat kidney cDNA library in lambda gt10 with PCR  
CC generated ET probes. These probes were also used to obtain the mouse  
CC ET gene (AAQ14340) from a mouse T-cell genomic library.  
CC An anti-sense ribonucleic acid molecule complementary to the  
CC indicated fragment in the features is also claimed.  
CC ET-1 is a bifunctional growth regulator, capable of stimulating  
CC the growth of some cell types while inhibiting the growth of others.  
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
CC in fact, antagonises this ET-1 activity.  
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
XX  
SQ Sequence 1779 BP; 304 A; 578 C; 547 G; 350 T; 0 other;

Query Match 69.7%; Score 747.2; DB 12; Length 1779;  
Best Local Similarity 98.3%; Pred. No. 1.5e-185;  
Matches 755; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
Qy 1 gggccccaccaggtgcccctggatggagagggccccagctcactcagctgcccagacca 60  
Db 1012 gggccccaccaggtgcccctggatggagagggccccagctcactcagctgcccagacca 1071  
Qy 61 caagcctgaagagagatgccctgtgataatgtcagcagctgtccctcccgatacc 120  
Db 1072 caagcctgaagagagatgccctgtgataatgtcagcagctgtccctcccgatacc 1131  
Qy 121 tgcctgccaactcacgtctggggagtggtgctgtgtccaatccagaggtgtctgtctgc 180  
Db 1132 tgcctgccaactcacgtctggggagtggtgctgtgtccaatccagaggtgtctgtctgc 1191  
Qy 181 tcggaccaccagcactgtgccccagcgatacacatgtgtgagctgaggggagtgatcag 240  
Db 1192 tcggaccaccagcactgtgccccagcgatacacatgtgtgagctgaggggagtgatcag 1251  
Qy 241 cgaggaagcgagatcgctgagctgaggaagatccctgcccgccgcttcttacc 300  
Db 1252 cgaggaagcgagatcgctgagctgaggaagatccctgcccgccgcttcttacc 1311





XX PD 17-OCT-1991.  
XX PF 03-APR-1991; 91WO-US02321.  
XX PR 13-MAR-1991; 91US-0083796.  
XX PR 03-APR-1990; 90US-0504508.  
XX PA (BRIM ) BRISTOL-MYERS SQUIB.  
XX PI Shoyab M, Plowman GD;  
XX DR WPI; 1991-325168/44.  
XX DR P-PSDB; AAR14325.  
XX  
PT New cysteine-rich growth modulating proteins, epithelins - useful  
PT as inhibitors of neoplastic cell growth and to promote wound  
PT healing and treat psoriasis  
XX  
PS Disclosure; Fig 18; 97pp; English.  
XX  
CC ET-1 and ET-2 were isolated from rat kidneys and their amino acid  
CC sequences determined. A full length rat ET cDNA was obtd. by screening  
CC a rat kidney cDNA library in lambda gt10 with PCR generated ET probes.  
CC These probes were also used to obtain the mouse ET gene (AAQ14340) from  
CC a mouse T-cell genomic library. ET DNA was also obtained from human  
CC sources (AAQ14339).  
CC An anti-sense ribonucleic acid molecule complementary to the  
CC indicated fragment in the features is also claimed.  
CC ET-1 is a bifunctional growth regulator, capable of stimulating  
CC the growth of some cell types while inhibiting the growth of others.  
CC ET-2 is functionally similar to ET-1 w.r.t. growth inhibitory  
CC bioactivity. In contrast, however, ET-2 is apparently not capable of  
CC eliciting the growth stimulatory activity characteristic of ET-1 and,  
CC in fact, antagonises this ET-1 activity.  
CC See also AAQ14338-40, AAQ14952-53, AAR14328-9 and AAR15315-20.  
XX  
SQ Sequence 1767 BP; 359 A; 499 C; 491 G; 418 T; 0 other;  
  
Query Match 43.8%; Score 469.2; DB 12; Length 1767;  
Best Local Similarity 76.7%; Pred. No. 7.3e-113; Indels 6; Gaps 1;  
Matches 589; Conservative 0; Mismatches 173;  
  
QY 1 gggcccccacagtgctgctgagagagcccccagctcactcagctccagcagccca 60  
DB 1006 ggaagcttcaggtaacctggaagaaagtcacggcctccctcagctccagcagccca 1065  
QY 61 caagccttgaaagagatgccccctgtgataatgtcagcagctgtccctcccgataacc 120  
DB 1066 cagatcttgaaagatgctccctgtgactcagctcagctgctccttaacaataacc 1125  
QY 121 tcttcccaactcagctggggagtgaggctgtctcccaatccagagctgtctgtgc 180  
DB 1126 tgcgtcagactcagctggggagtgaggctgtctcccaatccagagctgtctgtgc 1185  
QY 181 tcggaccacacagctgtgccccagcagatcacagctgtgtagtgaggggcagtgctcag 240  
DB 1186 ttgacacacacagctgtgccccagcagatcacagctgtgtagtgaggggcagtgctcag 1245  
QY 241 caggaagcagagatggtggctggagagagatgcctgcccgcgcgcgggttcctattacc 300  
DB 1246 aaggagacagatggtggctggagagagatgcctgcccgcgcgcgggttcctattacc 1305  
QY 301 caacccagacagatggtggctggagagagatgcctgcccgcgcgcgggttcctattacc 360  
DB 1306 caacatggagatggtggctggagagagatgcctgcccgcgcgcgggttcctattacc 1365  
QY 361 ccagaccaggggtggagctgggctgctgcccagtgcccagtgctgtgtgctgcagagat 420  
DB 1366 ccaagcctgaagggaagtggtggcctgctgcccagtgcccagtgctgtgtgctgcagagat 1425  
QY 421 cgcagcagctgctgcccggctggctacacacctgcaacgtgaaggctcgtcctgcgagag 480

DB 1426 cggcagcactgttgcgggctgggttacacctgcaacgtgaaaggcagaaacctgtgagaag 1485  
QY 481 gaagtgtctctgccagcagcctgcccacaccttctcctggccgtagccctcagctgggtgtgaag 540  
DB 1486 gatcagcgtctgtccagccttccatggacctgaccttggctctaaggttgg-----g 1539  
QY 541 gacgtgagtgagggaagagacacttctgccaatgataacacagacactgctgcgagagcaac 600  
DB 1540 aatgtggaatgtgtgcccggacatttctgccaatgataacacagcctgctgttgaagacagc 1599  
QY 601 cgacagggtggtgctgctgctccctacgcccagggcgctgtgtgtgctgctgctgctgctg 660  
DB 1600 caagagagctgggctgctgctccctatgtaaaggggtgtgctgctgtagagatgagcgtcac 1659  
QY 661 tgcgtcctgctggttccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 720  
DB 1660 tgcgtcctgctggttccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1719  
QY 721 cgcgctggtgagcggccttctgaggaccacccagccttgagacagctgctg 768  
DB 1720 cctcgtggacatacttctgaggatccagcccaagaccgctactg 1767  
  
RESULT 10  
AAQ14340  
ID AAQ14340 standard; DNA; 1767 BP.  
AC AAQ14340;  
XX 17-JAN-1992 (first entry)  
DE Mouse epithelin precursor.  
DE Mouse epithelin precursor.  
KW ET; anti-sense RNA; growth regulation; inhibition; stimulation; ss.  
XX Mus musculus.  
OS Mus musculus.  
FH Key 8..1774  
FT misc\_RNA  
FT Location/Qualifiers  
FT /\*tag= a  
FT /note= "Claim 29, page 56"  
XX  
PN WO9115510-A.  
XX 17-OCT-1991.  
XX 03-APR-1991; 91WO-US02321.  
XX 13-MAR-1991; 91US-0083796.  
XX 03-APR-1990; 90US-0504508.  
XX (BRIM ) BRISTOL-MYERS SQUIB.  
XX Shoyab M, Plowman GD;  
XX WPI; 1991-325168/44.  
XX P-PSDB; AAR14327.  
XX  
PT New cysteine-rich growth modulating proteins, epithelins - useful  
PT as inhibitors of neoplastic cell growth and to promote wound  
PT healing and treat psoriasis  
XX  
PS Disclosure; Fig 23; 97pp; English.  
XX  
CC ET-1 and ET-2 were isolated from rat kidneys and their amino acid  
CC sequences determined. A full length rat ET cDNA (AAQ14338) was obtd.  
CC by screening a rat kidney cDNA library in lambda gt10 with PCR  
CC generated ET probes. These probes were also used to obtain the mouse  
CC ET gene (AAQ14340) from a mouse T-cell genomic library.  
CC ET DNA was also obtained from human sources (AAQ14339).  
CC An anti-sense ribonucleic acid molecule complementary to the  
CC indicated fragment in the features is also claimed.







OS  
XX

**SQ** Sequence 336 BP; 70 A; 87 C; 106 G; 62 T; 11 other; 44

Query Match 18.5%; Score 197.8; DB 22; Length 336;

Best Local Similarity 97.18; Pred. No. 4e-42;

Matches	233	Conservative	0	Mismatches	3	Indels	4	Gaps	3
---------	-----	--------------	---	------------	---	--------	---	------	---

**Qy** 764 tgctgtgaggac-agtactgaagactctgcagccctcgggacccactcggagggtgcc 822

022

Db 239 TGCTGTGAGGGACAAGTACTGAAGACTTTGCAGCCCTCGGGACCCCACTCGGAGGGTGCC 180

[illegible]

Qy 823 ctctgctcaggcctccctagcacctccctccctaaacaaattctccctggacccattctga 882

[illegible]

Db 179 CTCGTGAGGCTCCCTAGCACCTCCCCCTAACCAAAATCTCCCTGGACCCCATTCCTGA 120

[illegible]

Search completed: September 5, 2002, 18:30:15  
Job time: 10547 sec

search completed: 30  
Job time: 10547 sec